National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

DONALD DAVIDSON:
PHILOSOPHY OF LANGUAGE WITHOUT LANGUAGES?

BY

STEPHEN H. K. JONES

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of MASTER OF ARTS.

DEPARTMENT OF PHILOSOPHY

Edmonton, Alberta
Spring, 1993

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-82144-2

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Stephen Howard Kirk Jones

TITLE OF THESIS: Donald Davidson: Philosophy of Language Without Languages?

DEGREE: Master of Arts

YEAR THIS DEGREE GRANTED: 1993

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Stephen H.K. Jones
362 Glengrove Ave. W.
Toronto, Ontario
M5N 1W5

April 16, 1993

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled DONALD DAVIDSON: PHILOSOPHY OF LANGUAGE WITHOUT LANGUAGES? submitted by STEPHEN HOWARD KIRK JONES in partial fulfilment of the requirements for the degree of MASTER OF ARTS.

Bernard Linsky

Dr. Bernard Linsky

Dr. F. Jeffry Pelletier

Bruce Hunter

Dr. Bruce Hunter

Dr. Uri Margolin

April 15, 1993

To my Grandmothers,

Amy Jones and Jean Ware,

who are both ninety.

ABSTRACT

The startling conclusion of Donald Davidson's 1986 paper "A Nice Derangement of Epitaphs" is that "there is no such thing as a language." This is a very astonishing pronouncement to come from someone who for over two decades was optimistic that a theory of meaning could be given for a language, in the form of a theory that defined a truth predicate for that language. But now, if there is no such thing as a language, then there is nothing for a such a theory to be about. And yet, Davidson persists in his conviction that the only way to talk about meaning is in terms of a truth theory. This led some commentators to wonder whether the recent development is consistent with Davidson's past programme.

In this thesis, I explain the argument of "A Nice Derangement of Epitaphs." I show how Davidson's position has shifted to become more like that of Paul Grice. Davidson has moved towards Grice's position, in that he now allows for a greater influence of intentions over meaning. But he has kept himself distinct from Grice by retaining his views on the metaphysics of intention, meaning and belief; and in doing so, has insured consistency with his past philosophy of radical interpretation.

ACKNOWLEDGEMENTS

My first debt to acknowledge is to my supervisor, Dr. Bernard Linsky. His patience, his gentle prodding, his understanding of the philosophy of language, and his demands for clarity in both writing and structure have made this thesis what it is.

I would also like to thank the other members of my committee--Dr. Jeff Pelletier and Dr. Bruce Hunter. Dr. Pelletier introduced me to the works of Paul Grice and meaning-nominalism, and got me thinking about some of the issues discussed herein. Dr. Hunter's unflagging enthusiasm, and the interest he took in my work, encouraged me greatly.

I would like to express my gratitude to Dr. Uri Margolin, my external reader from the Department of Comparative Literature, for his careful reading of my thesis, and for the many suggestions for changes that he made. My thesis is greatly improved, as a result.

I would like to acknowledge a special debt to Michael Hymers. I have benefitted from the many stimulating discussions we have had--on Davidson in particular, and on philosophy in general.

Finally, I would like to thank my family and friends, for their support.

Table of Contents

Introduction.....p. 1

Chapter One: "A Nice Derangement of Epitaphs" .....p. 12

Chapter Two: Davidson and Grice.....p. 41

Chapter Three: Reactions and Replies.....p. 60

Chapter Four: "All the King's Horses".....p. 82

Bibliography.....p. 100

Introduction

In Davidson's estimation, being interpretable is a matter of exhibiting patterns in one's behaviour. "Pattern in what is observed", he writes, "is central to the intelligibility of an agent's behaviour." (Davidson, "Structure and Content of Truth", 1990, p. 317) Davidson's position is that to have thoughts, or to have a language, depends on the ability of an interpreter to detect patterns in behaviour, to the extent that this pattern can be described in an explicit, systematic theory. He writes: "The possibility of understanding the speech or actions of an agent depends on the existence of a fundamentally rational pattern." (Davidson, 1990, p. 320) Davidson maintains that a failure to find any concrete patterns gives us warrant to conclude that there was no rational behaviour present in the first place. In linguistic behaviour, "nothing... could count as evidence that some form of activity could not be interpreted in our language that was not at the same time evidence that that form of activity was not speech behaviour."[1] Patterns are important, because they are essential to "our ability to understand actions as done for a reason." (Davidson, 1990, p. 317)

In Davidson's well-developed programme in the philosophy of language, the pattern that rationalizes speech

---

[1]Davidson, "On the Very Idea of a Conceptual Scheme", in Inquiries into Truth and Interpretation, 1984, p. 185.

behaviour is the pattern captured by a Tarski-style theory of truth. Such a theory gives a description of the semantic structure of the language being spoken. That is to say, it demonstrates the contribution that a word makes to all of the sentences in which it can appear, and it shows the logical relations between sentences. A Tarskian theory of truth for a language, if correct, entails theorems ('T-sentences') that supply the truth-conditions for the infinite number of potential sentences of that language. This stems from Davidson's holistic account of meaning, in which a word or sentence cannot have meaning unless it is a part of a larger language; there cannot be isolated cases of meaning. This idea is expressed in what I will call the Holistic Thesis. The Holistic Thesis states that "to interpret a particular utterance, it is necessary to construct a comprehensive theory for the interpretation of a potential infinity of utterances." (Davidson, "Belief and the Basis of Meaning", 1984, p. 148) To know that a particular T-sentence is true, one must know that the semantic pattern articulated by the theory "optimally fit[s] the evidence about sentences held true by native speakers." (Davidson, "Radical Interpretation", 1984, p. 139) The requirement that the theory optimally fits the data of the totality of sentences held true by the speaker, I will label the Holistic Constraint. The pattern of sentences held true will, if we take the speaker to be rational, "reflect the

semantics of the logical constants, [such that] it is possible to detect and interpret those constants." (Davidson, 1990, p. 319) The interpretation of the non-logical words is another matter, and requires more than observing the pattern of sentences held true. To interpret names and predicates requires the interpreter to look for the events or objects in the world that regularly cause their application.[2]

The goal of Davidsonian semantics was to write a theory of meaning that can be known to be true based on evidence plausibly available to an interpreter.[3] On Davidson's holistic approach, empirical content across the entire theory increases as the points at which the theory amasses evidence accumulate: "a strong theory weakly supported, but at enough points, may yield all the information we need about the atoms and molecules -- in this case, the words and sentences." (Davidson, 1984, p. 225) The axioms of a theory of meaning gain sharper focus as more theorems containing the words that the axioms govern are confirmed. Once again, the important point concerns pattern: "[P]attern is central to the theory's power to extract, from facts taken singly are relatively directly connected with what can be observed, facts of a more sophisticated kind. ... From the point of

---

[2]Davidson, "Epistemology Externalized" (unpublished manuscript) p. 3. See also Davidson (1990) pp. 320-1.

[3]See Davidson (1984) p. 131.

view of the theory, the sophisticated facts explain the simple, more observable ones, while the observable ones constitute the evidential base for testing or applying the theory." (Davidson, 1990, p. 317) The observable facts are the attitudes of holding sentences true, while the unobservable facts are the theoretical constructs of belief and meaning.

What is a pattern, one might ask, but a regularity known over time? In the essay "Communication and Convention", Davidson denied that regularity over time is essential to linguistic communication: "Perhaps some will feel inclined to make it a condition of calling an activity linguistic that there should be such regularity. I have my doubts...." (Davidson, 1984, p. 277) What is missing, in Davidson's estimation, from any picture of linguistic competence that makes essential appeal to regularities in order to explain meaning is the ability to handle novelty, either inadvertent or deliberate, in the way that we use words. Davidson brings to our attention that we are not regular in our speech, in the way assumed by the standard theories. According to Davidson, irregular language-use "happens all the time; in fact, if the conditions are generalized in a natural way, the phenomenon is ubiquitous." (Davidson, "A Nice Derangement of Epitaphs" 1986, p. 433) The examples of irregular language use that Davidson focuses on in the paper just cited are malapropisms. Davidson

considers the phenomenon of malapropisms to "threaten standard descriptions of linguistic competence." (Davidson, 1986, p. 437) In particular, he has in mind the thesis that to know a language, or to be able to understand a language, depends on previously learning and mastering a fixed set of rules or conventions. A theory that relies on past regularities cannot supply an interpretation of the meaning of a malaprop. But malapropisms are readily understood, so there must be more going on than any regularity theory of meaning is able to capture. The picture described above, of the Davidsonian programme of the 1970's, is equally paralysed by the introduction of novel uses of words, such as malapropisms.

I should comment at this point that there are two senses of novelty. A sentence can be novel to a listener owing to the fact that it assembles familiar words in an combination new to the listener, or it can be a novel sentence because it contains a word new to the listener. In the former case, a listener armed with a theory of truth for that language can recursively define the truth conditions for all sentences, including the novel one. The interpreter is able to "manufacture" new sentences to garner new evidence for the confirmation of the theory. A theory of meaning stands in a law-like relation to the language, and as such, is confirmed by its instances, and supports counterfactual conditionals about what the truth-conditions

of a particular sentence would be were it to be uttered.[4] In formulating a sentence in the object language that has an extension that is new to the evidential base, the field linguist is relying on the fact that her present theory entails a counterfactual "If the speaker were to utter this sentence, it would have such-and-such truth conditions." If the interpretation of this novel sentence is successful, then that particular sentence acts as further empirical evidence for the well-confirmedness of the truth theory in effect. Such empirical testability of theories of truth was a cornerstone of the Davidsonian programme in semantics. The programme thrived on the first sense of novelty.

In the latter case of novelty, however, where an altogether new term is introduced into the language, *e.g.*, a malaprop, a theory that relies on past regularities is paralysed. It seemed that Davidson's established position was a variation on the regularity theory, in that he counselled us to look to the objects or events in the world that *regularly* cause the application of a certain word. In the case of a malaprop, the idea of its having a regular cause has no obvious applicability. Because of this, Davidson is prompted to admit that the phenomenon of malapropisms threatens even his own established programme.[5]

---

[4]See Davidson, "Radical Interpretation" (1984) p. 174, and Davidson (1990) p. 310, and especially Davidson (1990) p. 313: "T-sentences ... have the form and function of natural laws."

[5]Davidson (1986) p. 437.

This second case of novelty is perhaps more aptly described as irregularity. Why should, one might ask, a theory of meaning have to account for irregular language use, when malapropisms and the like are clearly errors? Just about every other empirical theory whose theorems stand in a law-like relation to the evidence has some way of distinguishing between good and faulty data. Davidson is somehow driven by the conviction that malapropisms really mean something: they have a literal meaning, or what he calls "first meaning".

An early attempt to explain meaning, which could help illuminate how malapropisms mean something, was Grice's programme to explain the meaning of sentences in terms of the intentions the speaker held in uttering those sentences. From the intention on the part of the speaker to get something across to the listener, we arrive at the idea of speaker's meaning -- the meaning intended by a speaker on an occasion. Grice's programme was to then show how the idea of the meaning of a word or sentence is derivative from speaker's meaning. It appears as if malapropisms are prime examples of speaker's meaning: a malaprop means something because the speaker intended it to affect the hearer in a particular way, and also intended this to be accomplished by means of the hearer's recognition of that intention. Davidson himself appears to take a view more or less along this line. He writes: "A malaprop ...means what its promulgator intends it to mean." (Davidson, 1990, p. 310)

Is Davidson merely being Gricean in his treatment of malapropisms?

In order to answer this, we should look at a crucial passage in the paper on malapropisms: "[N]othing should be allowed to obliterate or even blur the distinction between speaker's meaning and literal meaning." (Davidson, 1986, p. 434) Whereas Grice takes a reductionist stance toward the distinction between speaker's meaning and literal meaning, arguing that one is a derivative of the other, from this quotation we learn that Davidson believes we must safeguard the distinction by clearly separating the two. The difference for Davidson between speaker's meaning and literal meaning comes down to the difference between what a speaker, on a given occasion, meant by uttering certain words, and what his or her words mean. Clearly, he sees a respectable role in a theory of meaning for each. He goes on: "In order to preserve the distinction we must... modify certain commonly accepted views about what it is to 'know a language', or about what a natural language is." (Davidson, 1986, p. 434) The way in which Davidson suggests we modify our views about what a natural language is, is to give up the idea that there are such things.

In Chapter One, I will examine in detail the arguments of "A Nice Derangement of Epitaphs", particularly as they relate to the conclusion that there is no such thing as a language. Part of the focus will be interpretative, aiming

to get clear what Davidson's argument is. I will also consider several objections that are "internal" to the argument, and point to ways in which they can be answered by Davidson.

Grice has similarly deflationary views when it comes to natural languages. He considers natural languages to be generalizations over what a group of speakers mean by certain sounds. Grice does not end the reductionism there, but further reduces what an individual means to particular intentions on the part of the speaker. I will label the move from individuals to intentions as 'reductionism', and use the term 'nominalism' for the move from natural languages to individuals. The reductionist stance is taken specifically towards the concept of meaning itself: Grice reduces meaning to particular intentions on the part of speakers. Grice's nominalism occurs at the level of natural languages: languages are abstractions resulting from the activities of individual speakers. I suggest, as the statement of my overall thesis, that Davidson shares with Grice this nominalistic attitude with respect to what a natural language is, though he does not share Grice's reductionistic stance with respect to meaning. Davidson is not a Gricean reductionist because being so entails that one can attribute isolated instances of meaning, is in direct conflict with Davidson's thesis of Meaning Holism. Grice's programme and those that it spawned are the topic of

discussion in chapter Two. Davidson's criticisms of Grice, as well as the points of similarity between the two, are also examined. I conclude that they are closer in certain respects than they might appear. Davidson's refinement of Grice's position was to take the intuition that malapropisms, if they mean anything, mean just what the speaker intended them to mean, and show how this could be preserved in a theory of meaning that demonstrated, in a way that Grice never could, the systematic structure of language.

Chapter Three will look at several reactions to Davidson's "Epitaphs" paper -- one positive reaction in defense of Davidson, and two negative ones which attack him. Bjørn Ramberg defends Davidson's pronouncement that languages don't exist, but I found his reading of Davidson to be misleading, and partly responsible for the way that some confused ideas have been taken in the literature as central to the debate. I show how Ramberg's interpretation misplaces the emphasis of Davidson's argument, and hence gives the wrong impression to those who would take up the attack against Davidson. These people include the linguist Alexander George, and the philosophers Dorit Bar-On and Mark Risjord. I show in Chapter Three how George's argument against Davidson rests on the logical fallacy of confusing the order of quantifiers. Bar-On and Risjord give two strong objections against Davidson, which focus on

Davidson's own Holistic Constraint and Independence Requirement. I show how these objections can be met by Davidson.

Chapter Four will pursue an issue already touched upon in this introduction: the problem of error. Malapropisms are *prima facie* mistakes; the most common reaction to one is correction. And yet Davidson wants to include them into a theory of meaning. To avoid the absurd consequence that one can never make a mistake in the way one uses words, Davidson must show how it is still possible, malapropisms notwithstanding, to be in true cognitive error about one's language.

Chapter 1

"A Nice Derangement of Epitaphs"

The startling conclusion of Davidson's essay "A Nice Derangement of Epitaphs" is that "there is no such thing as a language." (Davidson, 1986, p. 446) A common reaction to this proclamation of Davidson's is to say that by denying that languages exist, he has repudiated his career-long quest for a theory of meaning for a natural language, and shot himself in the philosophical foot. In this vein, Ian Hacking notes that the paper in which Davidson's claim occurs has an air of retraction with respect to his lifetime enterprise in the philosophy of language. (Hacking, "The Parody of Conversation", 1986, p. 448) Davidson himself admits: "[this recent result] threaten[s] standard descriptions of linguistic competence (including descriptions for which I am responsible)." (Davidson, 1986, p. 437) Much of Davidson's past writings in the philosophy of language have centred around the idea of a theory of truth for a natural language. Borrowing from Tarski the idea that one can define a predicate T for a language L (the predicate "T-in-L") that will be an extensionally adequate truth-predicate for that language, Davidson claims that the theory that constructs such a predicate will suffice for a theory of meaning for that language. Tarski could neither define a truth predicate for variable L --that is to say, he

could not define a general concept of truth that would work for all languages -- nor can he say what truth predicates defined for different languages have in common.[6] Despite this, it seemed natural to assume that Davidson was giving the outline for an empirical enterprise of constructing truth-theories for natural languages -- English, Finnish, Japanese, etc. -- and that the "L" in "True-in-L" was merely a place holder for those languages. But now, if Davidson is right, and there is no such thing as a language, then there nothing for the philosophy of language to theorize about. Dorit Bar-On and Mark Risjord have argued to this effect, claiming that Davidson has denied a theory of truth both its subject matter and empirical content. (Bar-On and Risjord, "Is There Such a Thing as a Language?" 1992, p. 163)

As Hacking rightly points out, though, Davidson is not committing "philosophical suicide." (Hacking, 1986, p. 448) His intention cannot be to retract his earlier themes: the target must be elsewhere. In order to understand what exactly Davidson is aiming at, we need to understand in what sense Davidson intended the denial that languages exist. A clue to what Davidson is up to is provided when he qualifies the conclusion that "there is no such thing as a language" with the almost parenthetical tag "not if a language is anything like what most linguists or philosophers have supposed." (Davidson, 1986, p. 446) What, in Davidson's

---

[6]See Davidson (1990) pp. 285 & 289 for a discussion.

estimation, do most philosophers and linguists consider a language to be?

Davidson outlines three principles commonly held amongst linguists and philosophers concerning what words mean: meaning is "systematic, shared, and prepared." (Davidson, 1986, p. 436) The first principle, "Meaning is systematic," aims to capture the way in which language is a structured phenomenon. It is a widely recognized feature of linguistic utterances that they have identifiable parts that recur in different combinations in other utterances. Longer sentences can be formed by concatenating simpler ones, such that the meaning of the longer expression depends on the meanings of its parts, and the way in which they were conjoined. Formulating a theory of meaning that captures this systematicity of language has typically involved a theory that has a finite number of axioms governing the interpretation of words, and a finite number of axioms for the rules of composition, which, using this finite base, recursively defines theorems that provide the interpretation of a potentially infinite number of sentences. Such a theory would demonstrate the logical relations of utterances, and the way in which words systematically contribute to the meaning of the utterances in which they occur. The holding of such a theory enables speakers "to interpret utterances on the basis of semantic properties of the parts, and the structure of the utterance." (Davidson,

1986, p. 436) Davidson's Tarski-style theory of truth for a language captures the sort of systematic structure in language sought after by the principle "meaning is systematic".

It is hard to specify exactly what is meant by the second thesis, "meaning is shared." The principle is widely held, but in name only. For instance, a Fregean theorist might claim that *senses* are shared amongst speakers of the same language. The sense of a term is the manner in which it presents its reference. For instance, the term "82nd Avenue", though referring to the same street as the term "Whyte Avenue", does so in a different manner. On Frege's theory of sense and reference, if two people attach different senses to the same proper name, they speak different languages.[7] The conception advanced by a Fregean theory is that sharing the same senses is necessary for sharing the same language. A theorist who stresses the social character of language might say that it is a *communal standard* governing the correct way to use and intepret words that is shared amongst speakers. The social thesis submits that language is essentially a normative affair: meaning in part has to do with the correct or incorrect application of words; using a word correctly presupposes the potential for being corrected, which in turn implies that there is a body

[7]See Dummett (1986) p. 462, where he places Davidson in the tradition of Frege. See also, Dummett (1978) pp. 424-5.

of other people, significant enough in number to embody the standards of correctness. On this view, the sharing of a language, and hence, the sharing of meaning, comes down to belonging to this significant body of people -- the linguistic community. And finally, another way of spelling out the slogan "meaning is shared" could be expressed in Davidsonian terms. As we saw above, the natural assumption about Davidson's early papers such as "Truth and Meaning" (1967) was that he was attempting to model speakers of a common natural language as having theoretical knowledge of a theory of truth for that language. Hacking offers a version of this assumption:

> When *J* and *K* talk to each other in common English, they share a Tarski-style theory of truth about that language (or may be modelled as sharing such a theory). *J* uses it to interpret *K* that way, and *K* interprets *J* that way. (Hacking, 1986, p. 448)

A speaker and hearer of the same language share in common one and the same truth-theory, which makes it possible for them to interpret each other's actual and potential utterances.

Davidson wants greater generality than just this: he wants the principle to accommodate communication that occurs across languages.[8] For instance, a speaker of German

---

[8]See Davidson (1986) p. 438, Davidson (1990) p. 311, and Davidson (1984) p. 157. "[T]hough communication by speech does not... require that any two speakers speak the same way, it does, of course, demand a fit between how speakers intend to be interpreted and how their interpreters understand them."

utters "Es regnet", and an English speaker understands her as saying "It is raining." What is held in common by these two agents? Davidson claims that what is shared in successful communication is a common understanding on the part of the speaker and the interpreter of the speaker's words.[9] But this seems as mysterious as ever. Davidson himself takes the second principle to mean: "for speaker and interpreter to communicate successfully and regularly, they must share a method of interpretation of the sort described [above -- that is, having a finite base, and recursively characterized]." (Davidson, 1986, p. 436; my emphasis) What is shared in successful communication is a common understanding on the part of the speaker and the interpreter of the speaker's words, and hence there must be a shared method that brings them both from the same utterances to the same understanding. One might reject this inference, and argue instead that the speaker and hearer might not need to share any theory at all, provided that they arrive at the same understanding of the speaker's words via different methods. Davidson's reply might be to say that such a coincidence would not ensure regular, successful communication. This suggests (but does not provide) some

---

(Davidson, 1990, p. 311)

[9]It might seem odd that Davidson is appealing at this point to what appear to be Fregean "thoughts" to explain the intuitive notion of "understanding the same thing" by an utterance.

sort of pragmatic definition of what it is for two people to understand each other.

The third principle, "meaning is prepared," Davidson understands as saying that "the systematic knowledge or competence of the speaker or interpreter is learned in advance of occasions of interpretation, and is conventional in character." (Davidson, 1986, p. 436) A convention is first and foremost a regularity in action. Moreover, a convention is an intentional, as opposed to accidental, regularity. Most importantly, it is a regularity that is known by its participants to occur, and that knowledge provides a good reason for each individual to act in accord with the regularity. Not only is the regularity known by each person to exist, but each person knows that everyone else knows: the regularity is a fact of mutual knowledge. Typically, this type of knowledge is described as *tacit* mutual knowledge, or, that there is a tacit agreement among the parties that some regularity ought to be perpetuated in the interests of coordination.[10] It can only be described as tacit knowledge, because the speaker might not be able to state what she knows about the applicable conventions, but her behaviour can be explained by characterizing her as if she knew these things. Speaking in terms of tacit agreement implies that the convention is settled in advance, which is the importance of "meaning is prepared."

---

[10]See Lewis (1975) for a representative example.

The picture that we get from these three principles combined is something like the following:

> each interpreter... comes to a successful linguistic exchange prepared with a 'theory' which constitutes his basic linguistic competence, and which he shares with those with whom he communicates. ... [E]ach party has such a shared theory and knows that the others share his theory, and knows that the others know he knows (etc.) (Davidson, 1986, p. 442)

This is the conception of language that Davidson attributes to philosophers and linguists when they assert that languages exist. What does this conception of language involve? The claim is that there is an objective fact that makes it the case whether someone is speaking a particular language or not: whether the speaker is intentionally following the conventions of that language. A theory of linguistic conventions proposes that there is a regularity amongst a group of speakers to attach a particular meaning to a particular utterance or word. Describing knowledge of a language in terms of linguistic conventions matches our intuitions that speakers of French understand the same thing by "le café est chaud", that speakers of other languages do not understand unless they have knowledge of the applicable conventions of French.[11] On the convention-based analysis of language, there is a 'standard interpretation' given by linguistic convention. Linguistic competence consists of

---

[11]See Ramberg (1989) pp. 100 & 103.

bringing one's intended interpretation in line with this standard interpretation. Failure to do so amounts to linguistic error. Merely intending to follow a convention does not make it the case that one actually follow it. Rather, following a convention is an objective social fact. The conventions of a language are discoverable in the community that speaks the language. In cases of idiosyncratic behaviour, it is not the case that the agent is following some other convention, held only by her, but that she is breaking an existing social convention. It is in this sense that a language is said to exist, independently and objectively. A language is something that an agent can be in a relation to, the relation of "-is a speaker of-", (or, conversely: "-is the language of-"); and a language is something that can be the object of true or false beliefs.

Davidson's arguments against the conventional nature of language are found in the 1981 paper "Communication and Convention."[12] The thesis of that paper is that conventions are not necessary for communication. Of the features of conventions described above, Davidson singles out as his primary focus of attention the requirement that a convention must exhibit regularities. A regularity, he argues, can only be a recurring pattern over time. What is it, he asks, that recurs over time to make linguistic

---

[12]Reprinted in Davidson (1984) pp. 265-80.

communication conventional? It can only be, he submits, a regularity in the interpretation of publicly identifiable patterns of noise or markings, such that the same sounds or marks are interpreted in the same way upon each successive encounter. It is this claim that Davidson doubts. He gives an embryonic version of the argument of "A Nice Derangement of Epitaphs" in the paper on conventions, when he writes:

> It is difficult to say exactly how speaker's and hearer's theories for interpreting the speaker's words must coincide. They must, of course, coincide *after* an utterance has been made, or communication is impaired. But unless they coincide in advance, the concepts of regularity and convention have no definite purchase. Yet agreement on what a speaker means by what he says can surely be achieved even though speaker and hearer have different advance theories as to how to interpret the speaker. (Davidson, "Communication and Convention", 1984, p. 278; italics in the original)

Hacking is quite right in observing that "A Nice Derangement of Epitaphs" is "a substantial generalization" of the thesis of "Communication and Conventions". (Hacking, 1986, p. 449)

Davidson intends to show, in "A Nice Derangement of Epitaphs", that the three principles outlined above -- that meaning is "systematic, shared, and prepared" -- are incompatible with the phenomenon of malapropisms. An example of a malaprop: "The dentist told me I have to wear a recliner on my teeth." This malaprop resulted from a confusion of like-sounding words, in this case "recliner" for "retainer". But this is also not essential. For

instance: "There wasn't a drop of wind" is a perfectly good malaprop.[13] This malaprop is the result of a confusion, but of cliches, not similar sounding words. The point is, there is no recipe for the creation, nor the understanding of a malaprop. A malaprop may be unintentional or deliberate. A deliberate malaprop might have the perlocutionary intention of adding some humour to the conversation. Humour is a pleasant quality to malaprops, but it is not essential. Malaprops, whether deliberate or not, whether the result of a confusion of similar words or some other gaff, are all recognized by some sort of degeneracy or absurdity that would result if they are interpreted in the standard way.

Malaprops interest Davidson for two reasons. First, malaprops are cases not covered by prior learning, so they threaten the picture outlined above of language as a shared structure known in advance to the communicating parties. Davidson groups the phenomenon of malaprops with other instances of novelty, including the interpretation of

---

[13]Nor do malaprops depend on any of the other principles that people have offered in conversations with me on the topic. For instance, one person offered the rule that all malaprops have the feature of there being the same number of syllables in the 'mistaken' word as the 'correct' word. This undoubtedly helps explain why the speaker made the mistake he or she did, but it is not an essential feature of malaprops. Part of Davidson's point is that there are no rules for understanding malaprops, other than the rough rules of thumb and maxims, such as the one provided by my acquaintance: "To find out what it could mean, look for a similar sounding word with the same number of syllables."

unfamiliar words and new proper names. Malaprops can be grouped with a family of linguistic phenomena, that includes the interpretation of Spoonerisms (e.g., "a blushing crow" instead of "a crushing blow"), slips of the tongue, or garbled and incomplete sentences. The second reason, then, why malaprops are interesting to Davidson is the fact that they unavoidably involve the notion of linguistic error. Common-sensically, one errs in language when the way one uses words fails to match standard usage. In uttering a malaprop, one is, strictly speaking, using a word incorrectly. Davidson is of the opinion, however, that this notion of correct usage is transparent and philosophically uninteresting: in uttering a malaprop, I would be "wrong about what a good dictionary would say, or what would be found by polling a pod of experts whose taste or training I trusted." (Davidson, 1986, p. 434) For Davidson, malapropisms are counter-examples to the standard conception of linguistic error. The notion of correct usage notwithstanding, most malaprops are nevertheless easily understood. In the above example, I understand the speaker as saying that she will have to wear a retainer on her teeth. "We want," concludes Davidson, "a deeper notion of what words, when spoken in context, mean; and like the shallow notion of correct usage, we want the deep concept to distinguish between what a speaker, on a given occasion, means, and what his words mean." (Davidson, 1986, p. 434)

It is the third principle -- that the abilities and competencies described in the above picture are the result of knowledge of conventions -- that Davidson claims cannot survive the considerations brought on by thinking about malaprops and other 'non-standard' uses of language. The first and second principles can only survive after much alteration.

To argue against the principle that meaning is something known in advance to both parties, as stipulated by the third principle, Davidson introduces what he calls 'prior' and 'passing' theories for both speakers and interpreters. A passing theory gives the literal, or what Davidson calls "first" meaning of an utterance. First meaning is that which the speaker intends the utterance to have. For interpretation to go through, all that is needed is a shared passing theory; that is, the interpretation intended by the speaker must be aligned with the actual interpretation used by the hearer. Davidson writes: "The asymptote of agreement and understanding is reached when passing theories coincide." (Davidson, 1986, p. 442)

In contrast to passing theories, Davidson also conceives of the notion of a <u>prior</u> theory. Prior theories are what the speaker and hearer bring to a conversation. The interpreter's prior theory is the theory, held in advance, for interpreting the speaker. It is what interpreter is equipped with, up until the moment a

particular utterance of the speaker is interpreted. It contains all that the interpreter knows about interpreting the speaker before communication between the speaker and the hearer begins, including knowledge of what linguistic regularities the speaker may have exhibited in the past. It is the interpreter's best guess about how to proceed in interpreting the speaker. As for the speaker, her prior theory is a reflection of how she expects to be interpreted: it is the speaker's "picture of the interpreter's readiness to interpret along certain lines." (Davidson, 1986, p. 442) The distinctions are captured in the following matrix:

	Prior	Passing
Speaker (S)	What S believes H's prior theory to be.	What S intends H's passing theory to be.
Hearer (H)	How H is prepared in advance to interpret S.	How H actually interprets S.

(adapted from Davidson, 1986, p. 442)

Interpreting malaprops falls under the jurisdiction of the passing theory, because, by hypothesis, no prior agreement or learning is involved. A theorem, or T-sentence, of a passing theory might look something like this: "'Their relationship is strictly plutonic' is $true_{in\text{-}L}$ if, and only if, their relationship is strictly platonic." The passing theory has the formal structure of any theory of truth, so it is suitable for the interpretation of a language. Its field of application is "vanishingly small"

(Davidson, 1986, p. 443), in that it is employed to handle a novel and possibly unintentional use of a word or expression, which might never be used in the same way again. The passing theory is "geared to the moment" (Davidson, 1986, p. 441) in that it provides the interpretation of all the words used on a particular occasion of successful communication: "Every deviation from ordinary usage, as long as it is agreed upon for the moment... is in the passing theory as a feature of what the words mean on that occasion." (Davidson, 1986, p. 442)

Let me turn at this point to examine how the two commentators on Davidson's paper -- Hacking and Michael Dummett -- understood the distinction between prior and passing theories. I believe that such a digression will illuminate the distinction in question. It is a matter primarily of how to interpret Davidson, but it will be important latter on.

Hacking takes the prior theories of the speaker and hearer to be the theories which they bring to a conversation; and he takes the passing theories to be those which evolve over the course of the conversation.[14] A pair of conversers each start with their best theory of the other's utterances, and quickly move on to a passing theory when the prior theory is found to be lacking. The picture that Hacking reconstructs is one in which "each pair of

---

[14]See Hacking (1986) p. 453, and see Dummett (1986) p. 459.

conversationalists evolves its own language." (Hacking, 1986, p. 449)

Dummett, however, contends that Hacking has not accurately reconstructed Davidson's position. Dummett supports this claim by pointing out that at one point Davidson speaks of an interpreter's prior theory undergoing modification, brought on by actions of the speaker.[15] Dummett concludes from this that "[the hearer] has, at every stage, both a prior and a passing theory, both being subject to continual revision." (Dummett, 1986, p. 459)

Dummett takes the important constrast between prior and passing theories to be between long-range and short-range theories. (Dummett, 1986, p. 460) For Dummett, the speaker's long-range theory is a picture of her expectation how, *in general*, her hearer is disposed to interpret her.[16] The speaker's short-range theory concerns how she intends her hearer to understand particular utterances she makes, without intending for the passing theory to in any way influence or modify how words are to be interpreted under his long-range theory.

---

[15]I think Dummett has Davidson (1986) p. 441 in mind. There, Davidson writes: "As the speaker speaks his piece the interpreter alters his theory, entering hypotheses about new names, altering the interpretation of new predicates, and revising past interpretations of particular utterances in the light of new evidence."

[16]To keep the pronouns straight, I shall deem the speaker to be "she", and the hearer to be "he." Dummett, for the same reason, labelled the speaker (S) and the hearer (H), but still managed, despite this effort, to run into pronoun trouble!

Dummett criticises Davidson for the way in which he draws the prior/passing distinction. His criticism focuses on the fact that the way in which the prior and passing theories of the speaker are differentiated is different from the way in which the distinction between the hearer's prior and passing theories in drawn. The difference between the two theories of the speaker is the contrast between her expectations and intentions, while the difference between the hearer's two theories is a matter of expectation and actual use. Dummett cautions: "the same principle of distinction ought to be used in both cases." (Dummett, 1986, p. 460) Dummett's parting shot is to say that Davidson is confused about his own prior/passing distinction, and that if he considers it to be important, he ought to "draw it with precision." (Dummett, 1986, p. 460)

Let me give my own thoughts on Dummett, making clear where I think he is right, and where I think he is wrong. Firstly, I find it hard to believe that Davidson expressed himself in any way other than what he intended. Dummett interprets the speaker's prior theory to be how she wants her hearer to *usually* understand certain words that she has uttered. Davidson would say that there is no such usual intention. For instance, he says "there is no such thing as how we expect, in the abstract, to be interpreted." (Davidson, 1986, p. 443)

Dummett is right, however, in saying that the deeper distinction that Davidson wants between what a speaker, on an occasion, means, and what his words mean is the difference between the speaker's prior and passing theories. Dummett's observation to the effect that the passing theory has no influence on the prior theory is also quite accurate.

How does the prior/passing distinction help toward the conclusion that there is no such thing as a language? Davidson's argument is found in the following passage:

> [W]hat interpreter and speaker share, to the extent that communication succeeds, is not learned and so is not a language governed by rules or conventions known to speaker and interpreter in advance; but what the speaker and interpreter know in advance is not (necessarily) shared, and so is not a language governed by shared rules or conventions. (Davidson, 1986, p. 445)

For languages to exist in the specified way, we need to find something that is simultaneously systematic, shared and prepared, that will guarantee linguistic communication. Having a shared prior theory fits all the requirements of being systematic, shared and prepared, but it will not always guarantee communication. A shared passing theory will, but it isn't prepared. A passing theory <u>is</u> a theory of a language, though in no ordinary sense of 'language,' in that, if correct, it has the formal apparatus to provide the interpretation of potentially any arbitrary sentence that the speaker makes. The holding of a passing theory cannot be what linguistic competence is all about, because a

passing theory is transitory, and hence unlearnable, and certainly not governed by conventions.

As hinted at above, it is the third requirement -- that meaning is contained in a theory that is learned or agreed upon beforehand, and then applied to cases, that causes all the problems. Thus, Davidson concludes that we "should try again to say how convention... is involved in language; or... we should give up the attempt to illustrate how we communicate by appeal to conventions." (Davidson, 1986, p. 446)

At this point, I would like to bring out some minor objections, which will aid exegesis of Davidson's argument, in an almost dialectic fashion. The first objection to the above argument, is to point out that passing theories do not give the appearance of being systematic. Dummett gives voice to this objection against the proposed systematic structure of a passing theory when he writes that the passing theory "will not be a structured theory, but only a collection of disconnected propositions." (Dummett, 1986, p. 466) Here, Dummett has in mind that the passing theory will be no more than a few 'meaning postulates', (*e.g.*, "'Epitaph' means epithet"), that is used in conjunction with, or to supplement, a systematic theory of meaning for a language. In order to preserve the systematicity of the passing theory, Davidson considers it necessary to say that when a malaprop is successfully interpreted, the deviant

word or words in the malaprop, when it appears at a location in the sentence normally occupied by the "correct" word or words, must "take over... the entire burden of that role, with all its implications for logical relations to other words, phrases, and sentences." (Davidson, 1986, p. 443) For instance, someone who correctly interprets "The dentist put a recliner on my teeth" must be prepared to give "recliner" all the logical properties in forming sentences that "retainer" had, or ordinarily has. The effects of interpreting "recliner" to mean *retainer* are felt "across the board" in the language. This is a direct consequence of what I have called Davidson's Holistic Thesis.

A further objection surfaces at this point: how does one get an infinite, holistic language out of a passing theory, since that theory provides the interpretation of a sentence spoken at an instant, containing a only finite number of words? To put the objection another way, from where do we get the material with which to interact in logical relations across an infinitary language? Davidson's answer is: "as treated in a prior theory, perhaps." (Davidson, 1986, p. 443) Davidson admits that malaprops can only be understood if I understand the other words appearing in the sentence with the deviant word. I wouldn't understand "The dentist put a recliner on my teeth" if I didn't understand the words "dentist," "put," and "teeth." In this case, Dummett is quite right to observe that the

passing theory "massively reduplicates" the prior one. (Dummett, 1986, p. 466)

It appears as if Davidson is dangerously close to the position that he is arguing against, which holds that knowing what language is being spoken is a crucial piece of knowledge in the process of interpretation. Davidson acknowledges that a prior theory is what most people would naturally consider to be a natural language. But, he claims that not even having a prior theory could be considered knowing a language. Prior theories are not languages because they are neither shared nor prepared. Davidson feels this is supported by two sets of considerations. Firstly, as he points out on several occasions, prior theories can differ from individual to individual, according to our varying expectations about the competence of the different people with whom we converse.[17] Davidson assumes that there can be different languages for different people, to the point that each person can speak a unique language, making the idea of a communal language normatively insignificant. Dummett astutely picks up on this fundamental assumption of "A Nice Derangement of Epitaphs":

> [Davidson's] approach does not entail repudiating that notion of a language

[17]For example, see Davidson (1986) p. 443: "one way to appreciate the difference between the prior theory and our ordinary idea of a person's language is to reflect on the fact that an interpreter must be expected to have quite different prior theories for different speakers." (emphasis added) See also ibid., p. 441.

> according to which English and Dutch are languages; but it involves taking the notion of an idiolect as fundamental to an account of what language is, and explaining a common language as a range of largely overlapping idiolects. (Dummett, 1986, p. 462)

An idiolect is an idiosyncratic language; a dialect spoken by one person. The second point that Davidson feels supports the idea that prior theories need not be shared at the outset of a conversation is to consider unintentional malaprops. A speaker of a malaprop expects to be interpreted using a theory that gives the meaning of "Recliner on the teeth" as *retainer on the teeth*. On the other hand, for someone who is in the position of interpreting this speaker for the first time, the prior theory would say that "Recliner on the teeth" means recliner on the teeth! So the prior theories are not shared. It could also happen, as is most likely in the case of a deliberate malaprop, that the speaker knows that the hearer is prepared in advance to interpret 'recliner' to mean recliner. This would mean that the prior theories <u>are</u> shared. In this case, the priors theories are shared, but this is not sufficient for successful communication if the speaker then goes on to say "recliner" to mean retainer. Davidson also suggests that the prior theories need not be the same when he writes that a speaker can act so as to cause the interpreter to modify the prior theory he holds for that speaker.

Let us now turn to how, in the light of malapropisms, Davidson thinks we ought to conceive the first two principles of meaning that are commonly held by linguists and philosophers. He feels that the first principle -- that meaning is systematic -- does not need to altered much in order to accommodate the phenomenon of malapropisms: it is a simple matter of insisting that passing theories are truth theories. They wouldn't provide an interpretation of a malapropism, Davidson asserts, if they did not have holistic structure, because to understand a sentence is to place it in the logical space of an entire language. In describing a passing theory as just a collection of disconnected meaning postulates, Dummett is focusing his attention too much on the T-sentences, and not enough on the truth-theory that entails them. A disconnected postulate is powerless on its own to provide any interpretation, unless it is able to enter into logical relations with other expressions and sentences. Only a full-blown truth theory has the power to do this. The fact that most of this 'material' with which to accomplish this comes most often from the prior theory is not an indictment of Davidson, but rather exactly what he had in mind.

In light of malaprops, Davidson feels that the second principle must be understood in an unusual way. We already see in the second principle Davidson turning away from the idea that speakers of the same language share objective

meanings, and placing the emphasis on sharing a theory. Davidson persists in calling the method of interpretation mentioned in explicating the second principle a "theory":

> I shall henceforth assume there is no harm in calling such a method a theory, as if the interpreter were using the theory we use to describe his competence. (Davidson, 1986, p. 438)

The instrumentalism that Davidson is articulating here is such that we need only consider the interpreter *as if* the theory is what he knows. No mention is made of what the interpreter actually knows. The strategy is to provide a rational reconstruction of the interpreter's actions. Davidson has no trouble with attributing to the rationalized interpreter knowledge of the theorems of the theory, and he often speaks in terms of such knowledge as being tacit knowledge.

What is shared between two conversing people is not a single theory of meaning for a particular language, because communication can occur across languages. What is shared is one theory per speaker, relativised to that speaker, held by both the speaker and the interpreter, that interprets the speaker's utterances. Davidson writes:

> The sharing comes down to this: the interpreter uses his theory to understand the speaker; the speaker uses the same (or an equivalent) theory to guide his speech. For the speaker, it is a theory about how the interpreter will interpret him. (Davidson, 1986, p. 438)

It appears from this as if what speaker and interpreter share is not the same theory. The theory used by the interpreter is "about" the meaning of the speaker's words. But the theory held by the speaker is "about" the interpreter's theory.[18] Davidson has to admit (as we see in the above quotation) that the two theories are mere *equivalents*, in some sense, of each other. But what is it for two theories to be equivalent theories? In what language is the comparison for equivalency made? Suppose, to continue the malaprop example employed above, in interpreting the speaker a hearer devises a T-theory for the speaker's language that contains the following theorem: "'Dentists put recliners on teeth' is $true_{in\text{-}L}$ iff dentists put retainers on teeth." This is unobjectionable, because the interpreter has no choice but to phrase the truth-conditions of the sentence in the object language in a meta-language which happens to be the interpreter's own language. But what about the speaker? If, according to Davidson, speaker and hearer share one and the same theory (or an equivalent one, understood in a certain way), then what does the theory look like when held by the speaker? What does the speaker know? If the malaprop is unintentional, one might say that it had better be tacit knowledge, if anything. But even still, it is in some ways absurd to

---

[18]This I take to be Dummett's point concerning "first order" and "second order" theories. See Dummett (1986) p. 466.

attribute to her knowledge (even tacit) of a theorem that contains the malaprop on the left-hand side of the bi-conditional, and the interpretation of the utterance, expressed in a language understood by the interpreter, on the right. For the speaker, "recliner on the teeth" means just what she thinks it means, namely "recliner on the teeth." Her T-theorem for herself would be "'Dentists put recliners on teeth' is $\text{true}_{\text{in-my-L}}$ iff dentists put recliners on teeth." Dummett considers this to be a problem. He writes: "it is indeed highly natural to say that [a speaker of a malaprop] believes that 'a nice derangement of epitaphs' means 'a nice arrangement of epithets'; but there is then a problem about how she represents this belief to herself, since she certainly does not represent it in those words." (Dummett, 1986, p. 468) This I will label the Representation Problem. The problem that Dummett puts to Davidson is how to describe what a speaker knows when she speaks a language.

There is an equivocation in the phrase "knowing a language". "Knowing a language" could imply the ability to speak a language, or the ability to comprehend or interpret utterances of the language. Davidson: "we usually think that having a language consists largely in being able to speak, but in [my theory] speaking will only play an indirect part." (Davidson, 1984, p. 157) Davidson's approach towards the problem of what it is to know a

language focuses solely on the interpreter, and what it is to understand a language. Dummett think that the Representation Problem is a genuine problem because he views the entire debate from the side of the speaker. For instance, he spends a page and a half "clearing up" Davidson's own "confusion" about the prior/passing distinction, but does so only from the point of view of the speaker. Dummett is interested in what the speaker must know, in order to be able to speak a language. Hacking points out that the bulk of Davidson's writing in the philosophy of language over the past 20 years is "entirely one sided: interpreter-sided." (Hacking, 1986, p. 448) This is an accurate assessment; and it applies to "A Nice Derangement of Epitaphs" just as well. For instance, Davidson writes: "...we lose nothing in the investigation of ...meaning if we concentrate on the knowledge or ability a hearer must have if he is to interpret a speaker." (Davidson, 1986, p. 436) Clearly, this is counter to Dummett's emphasis on what the speaker must know or master.

Davidson does not consider the Representation Problem to be something that requires an answer. He writes: "the theory describes under what conditions a sentence of a speaker is true, and so says nothing about what a speaker knows." (Davidson, 1990, p. 312) A truth theory describes the infinite number of things that an interpreter knows when he understands the speaker. But a connection with the

speaker must be made somewhere, for Davidson persists in speaking of the speaker and interpreter sharing something. That connection is made via the speaker's intentions. The theory describes what the interpreter knows about the relevant intentions on the part of the speaker. The theory gives the propositional content of the intentions that the speaker must have, if she is to mean anything by her utterances. The speaker of the malaprop intends an interpreter to understand the utterance "Recliner on my teeth" as *retainer on my teeth*. An intention is a propositional attitude, and as such, is not relative to any particular language. A propositional attitude is not entirely independent of language, for, as we will look into in the next chapter, Davidson believes that only linguistic creatures can have propositional attitudes such as intentions. But an intention, as a propositional attitude, remains the same regardless of what language it is expressed in. The intention "aller à la banque" is the same intention as the intention "go to the bank". Davidson, then, by expressing things in terms of the speaker's intentions, does not run into the absurdity noted above of having to attribute to the speaker knowledge of the truth conditions of her own utterances couched in terms of her hearer's (or anyone else's) language. On Davidson's instrumentalism, it does not matter what exactly the speaker knows, so long as

we are able to say that it is as if she intends "Retainer on my teeth" to mean retainer on my teeth.

In presenting this position, Davidson is relying heavily on Paul Grice, and the Gricean mechanism: a speaker intends her words to be interpreted in a certain way, and must intend so with an eye to how she believes her words will be interpreted; on the hearer's side, recognition of the speaker's intention drives interpretation.[19] If this is the shared method of interpretation used by both speaker and hearer, one wonders what has happened to the requirement that this shared method of interpretation is systematic, has a finite base, and is recursive in character. This issue will not be pursued here, but will be picked up in Chapter Three.

Of the "Epitaphs" paper, Hacking says: "the problems -- but not the conclusions -- of the paper [would] fit well into a book of essays [on] Paul Grice." (Hacking, 1989, p. 448) My thesis amounts to a disagreement with Hacking: Davidson's conclusions <u>do</u> fit nicely into Grice's programme. The topic for the next chapter is to make clear what Grice's programme is.

---

[19]See Grice (1957).

Chapter 2

Davidson and Grice

The purpose of this chapter is to explore the relationship between the programmes in the philosophy of language of Davidson and Grice. I will begin by expounding Grice's programme, as articulated by him. Second, we will examine how the initial starting point supplied by Grice has been expanded into a full- blown position, known as meaning-nominalism, by Jonathan Bennett. Third, we will look at Davidson's criticisms of Grice. What Davidson has to say about Grice is not all negative, and the Gricean elements that Davidson has picked up will be the fourth topic of discussion. And finally, I will discuss the overall similarities between Davidson and Grice, particularly on the issue of how to settle the difficult question of determining both what a speaker means and what she believes, without assuming either.

Grice's aim was to elucidate a particular concept of meaning -- what he calls "non-natural" meaning. Non-natural meaning is contrasted to "natural" meaning. Natural meaning occurs in cases where there is some natural connection between the sign and its signification: where the two exist together, or where one is a symptom of the other, such as in "Smoke means fire," or "Those clouds mean rain." Non-natural meaning strictly involves cases of persons meaning

something by an action; as in: "By doing x, she meant that P." Non-natural meaning is at work in all manner of actions, signals, marks, or vocal utterances. In non-natural meaning, there is no salient connection between the sign and its meaning: the word "fire" does not resemble fire in any way, nor does it have any natural connection with it. By 'natural connection between a sign and what it signifies, Grice's point is that something naturally meaning that *P* implies that *P*, whereas something non-naturally meaning that *P* does not imply that *P*.

The analysis of non-natural meaning pioneered by Grice is as follows: a person means something by a signal or utterance if she has the intention to produce some effect in her audience, or intends her audience to respond in a certain way, and intends this to be accomplished by way of the audience's recognition of her intention. (Grice, 1957, p. 76) To specify the effect intended by an utterance would be to give the content of its meaning. For instance, if the effect that a speaker intends by some utterance is for the audience to believe that the light is green, then the non-natural meaning of the utterance is "The light is green." Grice warns, however, that the content may not always be specifiable in a definite "that-clause" ("a belief that...").

Notice, however, that in articulating the content of the non-natural meaning of an utterance, we have quickly

moved from a description of the intended effect ("that the audience believe that P...") to a claim about the meaning of an utterance ("Utterance *X* means that *P*"). Mention of belief -- either the speaker's or the hearer's -- is elided from the meaning of the utterance. In other words, the meaning of an utterance that is not "that the audience believes that *P*", nor "that the speaker believes that *P*", but simply "*P*". Perhaps this objection is unfair: all that Grice wanted to explain was what it is to non-naturally mean something, not what it is to non-naturally mean that *P*. The only place where Grice (at least in his 1957 article, where he first put forward the thesis)[20] where he mentions that "X (non-naturally) means that *P*" is in connection with his vague remarks about conventions: viz., "what people (vague) intend ...to effect by X."

Grice recognises that the speaker might have several intentions in uttering a phrase, or there might be several effects that the speaker intends. For instance, I might say "The water is boiling" with the intention of getting you to realise that the water was boiling, and this with the intention for you to take it off the stove, and this also with the intention for you to make the tea. Now, did my uttering "The water is boiling" mean for you to make the tea? Grice states that it does not. What Grice is

---

[20]Grice's (1968) version of the thesis is that if I utter something meaning that *P*, I must intend my hearer to believe <u>that I believe</u> that *P*.

interested in is the primary intention of the utterance -- the one which brings about its effect only if it is recognised as having that intended effect. Call the initial intention a "Gricean intention." The claim, then, is that Gricean intentions are sufficient for meaning: meaning is a kind of intention. (Bennett, 1976, p. 11) Grice writes that though the effect of you making the tea depends on you realizing that the water is boiling, "it cannot be regarded as relevant to the meaning of my utterance." (Grice, 1957, p. 77) How can he make this claim?

Some theorists following Grice have attempted to rescue him from this difficulty. They claim that Grice's picture so far developed is inadequate. If the initial intention relevant to meaning is to get you to believe that the water is boiling, then the attempt at meaning is successful when, and only when, you believe that the water is boiling. But, they claim, I still mean something by "The water is boiling" regardless of whether or not you <u>actually come to believe</u> that the water is boiling. In other words, success at meaning something is not dependent on the intended effect actually occurring. The first intention, they claim, must be expressed in terms of <u>informing</u> the hearer that the water is boiling. And this intention, is successful when, and only when, you recognise my intention to inform you of something, whether you actually come to believe what I am attempting to inform you about, or not. This intention is

known in the literature as an *illocutionary* intention. It is successful when the utterance is understood as an attempt to inform the hearer of something, not when the hearer actually comes to believe what is being conveyed. It is this shift in theoretical emphasis, I believe, that is responsible for the opinion (held by Dummett) that there must be a convention in the language that labels certain sentences as assertoric. To intend to get you to do something, like make the tea, is a *perlocutionary* intention. It is successful when, and only when, you make the tea. On this theory, perlocutionary intentions are performed by way of illocutionary intentions. Illocutionary acts are entirely linguistic acts.

Grice explained non-natural meaning in terms of intentions. He hoped to explain linguistic meaning, or what he called "timeless" meaning, in terms of "what people (vague) mean by X" (Grice, 1957, p. 76) For this reason, Grice's approach, and those that it has spawned, are considered to be nominalistic. Nominalism is the broad philosophical position that universals do not exist; only particulars do. Universals are artifacts arising out of our way of speaking about individual things. Nominalists believe that talk of universals is extraneous, and that we can get along without them by speaking only of their particulars.

Grice's nominalism can be thought of along the same lines: it considers statements of the form "X means P in language L" to be artifacts of speaking of "what L-speakers mean by uttering X"; and those sorts of statements in turn to be artifacts of speaking about "what a particular person meant by uttering X". On this view, the concept of a language is not a primitive, but rather is something that should be somehow defined or generated from speaker-meaning. The notion of speaker-meaning stems from the plausible idea that when a speaker utters something, there is something that she is trying to "get across" to her audience. The speaker wants to inform the hearer that something is the case, and thus get the hearer to believe that something is the case. In cases of speaker-meaning, the speaker utters words that she believes will cause her audience to believe that something is the case. Speaker-meaning is contrasted to sentence- or word-meaning. Sentence-meaning is supposed to be what the sentence means in the language, independent of any individual's intention to use the sentence to mean something on a particular occasion. As ill defined as this idea is, it is *prima facie* plausible, in that, common-sensically, words and sentences have determinate meaning, whereas any speaker-meaning can be associated in the mind of the speaker with potentially any sentence. The distinction between speaker-meaning and sentence-meaning rests on intuitions that I can use words that ordinarily are

understood by hearers to mean one thing, while the speaker-meaning I have in mind is different. Normally, speaker-meaning and sentence-meaning coincide: I usually say what I mean. But occasionally, speaker-meaning and sentence-meaning diverge, often out of ignorance or error.

Meaning-nominalism, however, denies that sentence-meaning takes precedence in explaining meaning. True to its nominalist tendencies, it claims that one cannot make sense of what a word means in a language independent of how individual speakers of that language mean by their words. The first thesis of meaning-nominalism is: (1) "the concept of [sentence]-meaning can be elucidated through that of [speaker]-meaning, and this order cannot be satisfactorily reversed." (Bennett, 1973, p. 141) The thesis states that the meaning of an utterance on an occasion has a explanatory primacy in explaining facts about what speakers of a language generally mean by that utterance, or what that utterance means in a language, and so on. This is the central idea of meaning-nominalism: it "treats as basic the individual instances of meaning, by one speaker at one time, and gives a derivative status to every kind of general statement about meaning." (Bennett, 1976, p. 9)

How is the idea of meaning in a language to be generated in a nominalist programme? Obviously, "what X means in language L" is connected in some way with "what L-speakers mean by X", and that in turn with "what a

particular L-speaker meant by X on an occasion", but how? Generality, that is, claiming that "what L-speakers *generally* mean by X" provides the "meaning of X in L", is not sufficient, because it makes sentence-meaning a statistical matter, discoverable by perhaps taking a survey. Also, such an uncritical linguistic democracy seems unable to account for the phenomenology of linguistic error, in which divergence between speaker-meaning and sentence-meaning is considered to be wrong, and not merely unusual.

The direction that meaning-nominalism takes to tie individual speakers with meanings via an inter-personal standard is to employ the notion of a linguistic convention. Meaning-nominalists assume that linguistic meaning (meaning in a language) is nothing but conventional meaning. Bennett in particular deliberately conflates 'what an utterance means' with 'what speakers conventionally mean by that utterance'. For instance, Bennett writes: "the phrase 'conventional meaning' is pleonastic." (Bennett, 1973, p. 155) A convention is a regularity in actions across a group of people, intentionally conformed to by an individual member of that group because the fact that others conform to it gives her good reason to also do so. Conventions are able to satisfy the requirement for an objective, external fact, independent of the intentions of the speaker, about which she can be right or wrong in her beliefs, that is fundamental to sentence-meaning. Considering language to be

conventional seems to save the distinction between speaker-meaning and sentence-meaning: I can use words that possess a particular conventional sentence-meaning, but mean by them (in the speaker-meaning sense) something else quite different. Similarly, this way of speaking saves the phenomenon of linguistic error: I might believe that certain words have a conventional meaning that they in fact do not have.

However, as it stands at this juncture, meaning-nominalism has not completed its promised reduction. So far, we have made sense of the idea of what it is for L-speakers to mean by X that P -- when it is a conventional regularity that speakers mean P by X. But what it is for an individual speaker to mean P by X? Up to this point, it has been left unanalysed. This lacuna is made conspicuous when Bennett writes: "someone who utters something giving it a certain meaning need not conform to any convention for utterances of that kind." (Bennett, 1973, p. 141)

This prompts the articulation of the second thesis of meaning-nominalism: (2) "cases of meaning need not in any way involve conventional meaning"; (Bennett, 1973, p. 141) What the second thesis is implying is that there is a more basic concept of "meaning-in-general," prior to that of conventional or linguistic meaning. The overall strategy of meaning-nominalism, then, is to characterize this concept of meaning as broadly as possible, and from there, one can

explain linguistic meaning as the intersection of meaning and conventions. (Bennett, 1976, p. 25)

The broad base of meaning-in-general on which to build their theory is non-natural meaning, explained above. Most occasions of non-natural meaning are un-structured signs, signals, or actions, usually with a strong iconic element. For instance, I might pantomime that the fish are biting. Meaning-nominalists begin their analysis of meaning there, with utterances that are assumed to lack grammatical structure. (Bennett, 1976, p. 7) Conventions, it is argued, arise out of the wide class of non-natural meanings. As tokens of non-natural meanings get repeated in recognisable situations, their iconic, signalling qualities gradually drop out, leaving the convention standing more or less on its own. It is only at a much latter stage in the overall story told by meaning-nominalism that the notion of a meaningful utterance part emerges.

Meaning nominalism wants to derive linguistic, conventional meaning from non-linguistic, non-conventional meaning, without assuming anything about language. To do this, it must be able to prove that one can have grounds for attributing intentions solely on the basis of non-linguistic behaviour. To do otherwise would beg the question. Broadly, the plan is to start with thought first, and then proceed via meaning-in-general and conventions to the concept of conventional meaning. Bennett seems confident

that close observation of behaviour can support the ascription of the kinds of beliefs and intentions demanded by the theory:

> To develop a story along those lines we should need ...a steady hand for the management of subtle and complex behavioural details; but no doubt it could be managed somehow. (Bennett, 1973, p. 148)

The Gricean intentions demanded by the theory are fine-grained, self-referring intentions to produce beliefs. By fine-grainedness, I mean that I can have the intention to walk along Whyte Ave without intending to walk along 82nd Ave, even though Whyte Ave is 82nd Ave. It is of the essence of language that if a speaker is to mean something, we must be able to finely discriminate meanings, such that she might mean P and not Q, even though P and Q stand or fall together.

Bennett recognises these difficulties that stand in the way of attributing non-linguistic intentions, and advocates a kind of parsimony in ascribing content to intentions: "the ascribed belief or intention must be as simple, as low-level, as contentless as the data permits; so that, for instance, we ought not to say 'It intended to make me afraid' if the behaviour is as well accounted for by 'It intended to make me run away'." (Bennett, 1973, p. 146) This suggests a kind of instrumentalism, whereby the content of intentions and beliefs is attributed only insofar as it explains and predicts behaviour.

In summary, meaning-nominalism can be characterized by the following points:

(1) conventional meaning is to be explained in terms of occasion-meaning (the meaning of an utterance on an occasion), and not the other way around;
(2) meaning is not necessarily conventional;
(3) language is essentially conventional;
(4) not all meaning is linguistic;
(5) meaning can be explained without the concept of language;
(6) meaning is a kind of intending;
(7) complex beliefs and intentions without language are possible.

Davidson is critical of Grice's overall approach, arguing that one cannot resolve beliefs and intentions in advance, and then solve for meaning. He writes:

> There is a principled, and not merely a practical, obstacle to verifying the existence of detailed, general and abstract non-linguistic beliefs and intentions, while being unable to tell what a speaker's words mean. ...[W]e have no good idea how to set about authenticating the existence of such attitudes when communication is not possible. ... If this is so, then an inventory of a speaker's sophisticated beliefs and intentions cannot be the evidence for the truth of a theory for interpreting his speech behaviour. (Davidson, 1984, p. 143-4)

As we can see from this, Davidson is adamant in his "refusal to get belief under way independently of meaning." (Bennett, 1976, p. 269) Davidson contends that as evidence for a theory of the interpretation of speech, beliefs and intentions are not sufficiently removed from the final goal, which is an articulation of the meaning of sentences. In his view, the three basic intensional notions of belief,

meaning and desire "occupy the same conceptual orbit".[21] In order to not beg the question by assuming in advance the concepts to be elucidated, the evidence for a theory of interpretation must be non-semantic and non-linguistic. With this, of course, Grice and Bennett would agree. The difference of opinion, then, comes down to the fact that Davidson considers descriptions of detailed beliefs and desires to be inherently semantic.

What is his argument for this? As Jonathan Bennett pointed out, much hinges on the word 'detailed' in the above quotation.[22] Only in language can one make fine-grained divisions between intensional states -- propositional attitudes, such as believing in P but not in Q, even though P and Q stand or fall together. For instance, I might know that a certain store is on Whyte Ave, but not know that it is on 82nd Ave, even though Whyte Ave is 82nd Ave. To individuate between states of my mind more finely than just to say that I am thinking about "that street out there", such that I can be thinking about 82nd Ave and not be thinking about Whyte Ave, seems only possible when one attributes to me a language that distinguishes between 82nd Ave and Whyte Ave. It is hard to see how one can attribute fine-grained intentions to a creature that lacks language: perhaps it is unobjectionable to say that my dog <u>wants</u> to

---

[21]See Davidson (1990) p. 315.

[22]Bennett (1976) p. 271.

greet the man in the yellow hat, but we cannot say that it wants to greet the man in the yellow hat but not greet the spy, even though the man in the yellow hat is the spy. Furthermore, Davidson points out, the attribution of beliefs and desires on the basis of solely non-linguistic evidence is vastly underdetermined.[23] For example, one might observe a person faced with choosing between an apple and an orange who chooses the orange. This might be the result of a preference for what is orange rather than red, for what appears the riper of the two, for what is juicier, or even for what is on the left rather than on the right. We have no clear way of settling the matter of a person's preferences apart from the interpretation of speech.

Davidson himself admits that these sort of considerations do not constitute a real *argument*.[24] Perhaps a better argument for showing that the ascription of complex beliefs and desires independent of the interpretation of words is impossible would be the following: (1) The sole evidence that is available to an interpreter would be the attitude on the part of the speaker of holding a sentence to be true. (2) A speaker holds a sentence true based on two factors: what the speaker takes

---

[23]Davidson (1984) p. 163.

[24]See Davidson (1984) p. 164: "These considerations will probably be less persuasive to dog lovers than to others, but in any case they do not constitute an argument." (underscoring added)

the sentence to mean; and what the speaker believes to be the case.[25] Since, from (2), meaning and belief are interdependent, (3) the interpretation of meaning can only proceed hand in hand with the attribution of beliefs.[26]

To answer the objection that Davidson has not found non-semantic, non-linguistic evidence for his theory of meaning, Davidson's reply is that the attitude of holding a sentence to be true is indeed intentional, but it does not require the individuation of intentional states. One can know of the existence of this attitude without knowing anything about the content of the sentence held true -- without knowing what its propositional object is.[27] Ramberg writes: "It is necessary to assume the attitude of holding a sentence to be true on the part of the speaker, but this assumption tells us nothing of intentions and beliefs of a kind useful in determining meaning." (Ramberg, 1989, p. 69) The assumption is not semantic, though it does presuppose that the speaker has a minimal capacity for semantic understanding. But the presupposition that the agent has a minimal capacity for semantic understanding

---

[25]See Davidson (1984) p. 142 and Davidson (1990) p. 318.

[26]Bennett is unimpressed with this argument. He writes: "[Davidson] has much to say about the attempt to distil out separate belief and meaning components from the total import of a linguistic utterance; but that is a world away from the attempt to attribute beliefs where there is no language. General slogans such as 'Meaning and belief are interlocked' tends to blur that vital distinction." (Bennett, 1976, p. 271)

[27]Davidson (1990) p. 323.

amounts to assuming that the agent has linguistic capacities. So, as an argument to show that belief and desire are only possible in linguistic creatures, it fails. It fails, as I have reconstructed it, because it assumes at the outset that the initial evidence for the theory is the holding of a sentence to be true. While it might be true that Davidson has assured himself a non-semantic base of evidence for the theory (*non-individuated* semantic states, at least), he has not found a non-linguistic starting point.[28] Davidson concedes this objection: "We set out to find an argument to show that only creatures with speech have thoughts. What has just been outlined is not an argument, but a proposal." (Davidson, 1984, p. 167)

Rather than argue, Davidson *proposes* that belief and meaning are two interlocking concepts that together explain the pattern made by the various sentences a speaker holds to be true.[29] This is done on analogy with Bayesian decision theory, where subjective utility and subjective probability together explain preferences amongst courses of action.[30] Given that one knows before hand the agent's preferences over actions or states of affairs, if one knew the value

---

[28]A similar objection can be found in Bennett (1976) p. 271.

[29]More recently, Davidson has refined this somewhat. Rather than relying only on sentences that are held true or false, Davidson now speaks of basing a theory of meaning on the degrees to which sentences are held true. See his (1990) Section III.

[30]Davidson develops the analogy with decision theory in (1984) pp. 145-8 and pp. 160-2, and also (1990) pp. 316-26.

that the agent places on a particular event happening, then one could solve for the degree of belief, or the likelihood that the agent deems the event will occur. By symmetry, if one knew the degree of belief, one could solve for the relative values of the agent. To complete the analogy, if one knew the meaning of a sentence, then one is well on the way towards attributing a belief to a speaker (what someone says being *prima facie* evidence for what they believe); and if one knew what belief a sentence expresses, one has a key to interpreting its meaning.

Rather than pursuing Grice's strategy of first determining what a speaker believes and intends, and then determining what the speaker means, Davidson opts instead for an approach that simultaneously determining both the meaning of an expression and the belief that it expresses.[31] An account of the meaning of an expression is given by a holistic theory of meaning, that, together with the complementary and similarly holistic theory of the speaker's beliefs, form a unified theory of the person and her actions.[32]

Davidson must somehow break into the circle (of his own creation) of the interdependence between belief and meaning. What he needs is a device that separates out the contributions of belief and meaning. The Principle of

---

[31]See Davidson (1984) p. 144 and Davidson (1990) p. 316.

[32]See Davidson (1984) p. 154 and Davidson (1990) p. 322.

Charity (P of C), Davidson proposes, is just that device. I will not launch into a protracted discussion of the pros and cons of the P of C here. In essence, it allows an interpreter to "hold belief constant and solve for meaning." (Davidson, 1984, p. 137)[33] So, the way that Davidson breaks into the circle of interdependence formed between meaning and belief is via belief. The methodology pursued by a radical interpreter, using the P of C as a presupposition, is to "assign to sentences of a speaker conditions of truth that actually obtain (in our opinion) just when the speaker holds those sentences true." (Davidson, 1984, p. 196) What this amounts to is an assumption that the speaker's belief is just the same as the belief that the interpreter would form under the same circumstances. Though assumptions about beliefs have a certain methodological priority, this doesn't change the fact that meaning and belief are simultaneously attributed, with alterations in one resulting in necessary changes in the other.

Ramberg asserts that the initial stage of radical interpretation, which selects the intentional state of holding a sentence true as the evidence for a theory of meaning, "[is] not sneaking 'meaning' in through our

---

[33]See also Davidson (1990) p. 319: "Hold one factor steady ...while determining the other"; and Davidson (1984) p. 167: "[W]hat is needed is a method for holding one factor steady while the other is studied."

assumptions about beliefs and intentions." (Ramberg, 1989, p. 69) However, the same cannot be said about the next step in the radical interpretation methodology, which begins by assigning specific content to those sentences held true. As mentioned above, the P of C is the device that separates the contribution made by the speaker's beliefs and what the speaker's words mean. I argue that to assume, as the P of C does, that when a sentence is held true the belief it expresses is also true does sneak meaning in through assumptions about beliefs. I do not, though, intend this as an indictment of Davidson. Rather, I raise this to point out the similarity between Davidson and Grice. I suggest that they are much closer than they appear. Against Grice, Davidson asserts that fine-grained beliefs and intentions cannot be individuated independently of the interpretation of speech. But loosely following Grice, Davidson's plan is to attribute meaning based on assumptions about beliefs and intentions.

Chapter 3

Reactions and Replies

To review, I argued in the last chapter that Davidson is not a Gricean because of what he considers to be an impossible requirement of that programme: to individuate in advance intentions finely enough to support the attribution of meanings. With Grice, he believes that communal languages are nothing but the aggregate of the way that individuals speak. Unlike Grice, however, he does not believe that what individuals mean reduces to anything more basic, such as intentions. Rather, beliefs and intentions go together with meanings to form a unified theory of the individual's speech behaviour.

In this chapter, we will look at Ramberg's Davidsonesque arguments for the thesis that there is no such thing as a language. I will show where it falls short. After that, we will look at two criticisms of Davidson by George and by Bar-On and Risjord -- criticisms that, I must admit, are partly directed at Ramberg's misinterpretation.

As we saw in chapter One, Davidson's argument for why there is no such thing as a language (at least, not in the way that philosophers and linguists think that languages exist) is because there is nothing that is systematic, and most importantly, both shared and prepared, that is adequate, in all situations, for successful communication.

The commonly held conception of a language is that it is a structural regularity, expressed in terms of a set of rules or conventions, which is learned or acquired by mastering in advance those rules or conventions, the knowledge of which is applied to particular cases. It is in virtue of common possession of a language as such an abstract entity that linguistic communication between two people is said to be possible. Davidson considers the apparatus of prior and passing theories to be fatal to this description of what a language is because we quickly realize that nothing fits the bill. What we need is something that is shared and learned in advance ("prepared"), which is adequate to interpretation, or makes communication possible. On Davidson's prior/passing distinction, what is learned in advance is the prior theory. But the prior theory is not always sufficient for communication: its incompleteness is exposed by a single malaprop. By this, I mean that the prior theory is unable to provide the correct interpretation of a malaprop. When someone tells us to "Remember Pearl Island!" a prior theory would probably inform us that the speaker is urging us to remember Pearl Island. But this is obviously not right. What is shared, on the other hand, is the passing theory, but the idea of learning a passing theory has no applicability, because a passing theory is constructed on the spur of the moment. So, there is nothing that is both shared and prepared that is sufficient for

successful communication. This result precipitates Davidson's conclusion that "there is no such thing as a language, not if a language is anything like what many philosophers and linguists have supposed." (Davidson, 1986, p. 446)

Ramberg takes Davidson to be saying in the denial that the prior theory is not *necessary*, either. He reconstructs Davidson's argument as saying that there is no such thing as a language, at least not in the way that most philosophers of language and linguists think it exists -- namely, as incorporated in rules or conventions -- because knowledge of those rules or conventions is neither sufficient nor necessary for communication to occur. "It would not be necessary," writes Ramberg, "since radical interpretation proceeds without such knowledge."[34] What Ramberg has in mind is the fact that a 'field linguist' in a circumstance of radical interpretation, having no prior knowledge of the language of her subjects, not knowing in advance how to begin assigning truth conditions to the speaker's sentences, can nevertheless construct a theory of truth "from scratch," as it were. It follows from this, according to Ramberg, that a prior theory is never necessary. Ramberg goes on to claim that Davidson's malaprop argument amounts to saying that if we used "*nothing but* malapropisms, communication

---

[34]Bjørn Ramberg, Donald Davidson's Philosophy of Language: An Introduction (Basil Blackwell: 1989) p. 104

would still be possible." (Ramberg, 1989, p. 101; emphasis in original)

I do not know how Ramberg regards the idea that we could communicate in nothing but malaprops to be implied by anything Davidson has said. Unfortunately, it has been accepted outright by some commentators as an "official" Davidsonian doctrine.[35] Critics have also picked up from Ramberg the notion that Davidson's denial that there are such things as languages amounts to the claim that there is nothing shared and prepared that is both necessary and sufficient for successful communication. For instance, we find: "[E]ven if the malapropism phenomenon helps show that the assumption of a shared, stable set of linguistic conventions does not *suffice* to explain what goes on in linguistic communication it does not show that the assumption is not a necessary one. The Davidsonian rejection of language, however, requires both claims." (Bar-On and Risjord, 1992, p. 186, n. 30)

There is ample textual evidence that Davidson's intentions are nothing like what Ramberg reconstructs them to be. I take Davidson all along to be concerned with what sort of theoretical description of linguistic competence, the propositional knowledge of which would be *sufficient* for understanding a language. For example, we read: "All that

---

[35]For instance, see Dorit Bar-On and Mark Risjord "Is There Such a Thing as a Language?" *Canadian Journal of Philosophy* (April, 1992) p. 185.

we should require of a theory of truth for a speaker is that it be such that *if* an interpreter had explicit propositional knowledge of the theory, he would know the truth conditions of utterances of the speaker." (Davidson, 1990, p. 312; emphasis in original) And also: "A theory of truth for a speaker is a theory of meaning in ... that explicit knowledge of the theory would suffice for understanding the utterances of that speaker." (Davidson, 1990, p. 312)[36] Davidson is not attempting to give necessary and sufficient conditions for knowing a language: his 'as if' instrumentalist stance only requires a description of an ability that is sufficient for communication, were a speaker or hearer to possess this ability. There is no claim as to what mechanism is actually going on.

Davidson considers prior theories to be more important that Ramberg makes them out. For instance, Davidson writes: "of course things previously learned [are] essential to arriving at a passing theory." (Davidson, 1986, p. 443) Ramberg seems to be confusing in what sense a radical interpreter does not have a prior theory. While it is true that initially she comes to the exercise with no prior knowledge of the language, it is wrong to say that she does not, at any time, have a prior theory, and conclude from this that prior theories are therefore not necessary. A

---

[36]For more on this theme, see Davidson (1984), essay 9: "Radical Interpretation."

radical interpreter can only collect data over time. When the evidential base increases, a prior theory results. An unavoidable assumption on the part of the radical interpreter is that the axioms that she collected yesterday are still good today, unless some recalcitrant sentence forces her to change the theory that she currently holds. So, the example of the radical interpreter cannot show what Ramberg wanted it to show, namely that having a prior theory is not necessary for successful communication to take place.

Ramberg gave the impression that the skill of an idealized radical interpreter nothing more than being able to continually come up with the right passing theory for interpreting single utterances. He argues that, in principle, it is possible for a radical interpreter to always come up with the correct theory, as the need for one arises, no matter how abrupt the switch in what language is being spoken is. Ramberg's position seems to suggest, in Dummett's words, that "[i]t would make no difference what sounds were spoken, as long as the intention behind them remained constant." (Dummett, 1986, p. 474) Thus, Ramberg concludes, if we spoke in nothing but malapropisms, communication would nevertheless not be hindered.

However, I would like to argue that the idea of speaking in nothing but malapropisms is incoherent. Malapropisms are only malapropisms when there is a salient contrast with 'normal' words. If one talks in nothing but

malapropisms, this contrast is lost. To speak in this way would be to continually change the way one uses every word in one's repertoire. Ramberg suggests that were someone to talk in this way, a radical interpreter could in principle always understand him or her. But this would require the interpreter not merely to rely upon wit and wisdom, but also be telepathic. The interpreter would have to keep up and match every change that the speaker is making to every word in the language, at every moment. Against Ramberg, I suggest that we cannot make sense even of the *idea* of someone who speaks in nothing but malapropisms. We cannot make sense of this idea because we cannot make sense of such a *speaker*.[37] We cannot make sense of someone continually erecting a language, and then at the next moment scrapping it altogether, and replacing it with an entirely new one. To speak in nothing but malapropisms would be to speak in a way such that everything one says breaks the pattern of consistent, stable behaviour; which is to say, no rational pattern can be found. As mentioned at the opening paragraph of this chapter, Davidson argues that in such a case one has warrant to doubt whether it is linguistic behaviour after all.

---

[37]I intend this to echo some of Davidson's more famous arguments, in particular, the argument that we cannot make sense of the 'very idea of a conceptual scheme' because we could not make sense of an "inhabitant" of a radically divergent conceptual scheme -- an argument that I alluded to at the beginning of the Introduction.

Ramberg has given the impression that what is important to linguistic communication is just coming up with the right passing theory to interpret single utterances from moment to moment.[38] In the course of this thesis, I have been putting forward the view that Davidson has shifted in his position towards nominalism. If the impression that Ramberg has given is correct, then no doubt Davidson has gone entirely over to nominalism, where individual utterances need not have any connection with any others. However, I contend that Davidson is adamant in retaining his holism. What is important, for Davidson, is not the ability just to come up with a successful passing theory, nor even to come up with successive passing theories, but the ability to take a prior theory, and revise and improve it, along holistic lines, in the face of new evidence, creating a passing theory where required. It is coming up with a succession of theories -- prior or passing -- that is the mark of linguistic competence. One of the obstacles facing a radical interpreter is deciding whether a sentence that doesn't fit the presently held prior theory is actually reason to modify the prior theory, or whether the anomaly can be allocated to a passing theory, where it does not

---

[38]For instance, Bar-On and Risjord write: "[In Davidsonian Semantics without language] [l]inguistic communication can now presumably be seen to be possible as long as -- and insofar as -- interpreters are able to construct passing theories designed to interpret single, individual utterances of speakers (cf. Ramberg, 106ff.)" (Bar-On and Risjord, 1992, p. 188)

necessarily force a change in the prior theory. To create a passing theory is to decide that the sentence has only a momentary, though still literal, meaning, and that this irregularity is not one that should perturb one's confidence in one's grasp of the language. Of course, if the malaprop is encountered enough times, the interpreter might decide that this is something that requires a change in the prior theory. In offering this explanation, I am relying on Dummett's evaluation of the distinction between prior and passing theories to be the difference between long-range and short-range theories.[39] I am also relying on the idea (also Dummett's) that the prior theory is the theory of what the speaker's words mean, and the passing theory is a theory of what the person meant, on that occasion.[40]

Ramberg is perhaps correct in pointing out the connection between radical interpretation and the interpretation of malapropisms, in that the skills that a radical interpreter relies upon are fundamentally no different than the skills that someone uses when they create a passing theory; both rely on "wit, luck, and wisdom." (Davidson, 1986, p. 446) Bar-On and Risjord argue bitterly against Ramberg about whether the knowledge that a radical interpreter has is always incomplete. I will not venture my views on this, except to say that the phenomenon of

---

[39]Dummett (1986) p. 460.

[40]See p. 29, above.

malapropisms suggests that whatever knowledge we glean, it is always possible that it be insufficient on some occasions.

Though Davidson can only show the insufficiency of prior learning of regularities, that is all he needs for his argument to be carried. His intention is to show that there is nothing that is learned in advance that is sufficient in all cases for communication to be successful. Having a shared prior theory -- that is, when the speaker is aware of exactly what way the interpreter is prepared to interpret the speaker -- is not (always) sufficient for communication to occur on all occasions. Why? Well, take for example a deliberate malaprop. The speaker knows in what way her audience in the past has interpreted her, and she knows that the method of interpretation they used in the past does not provide for the way she intends to be interpreted on the occasion, but she knows that her hearer will probably come to the intended interpretation, anyway. This situation surely is possible, and as Davidson says, it happens all the time. Since it is possible to come to an understanding of an unanticipated malaprop, it is not required for communication that the theory that an interpreter is equipped with at the outset of the exchange be one that the speaker intends him or her to use.

Alexander George believes that Davidson's premise is not enough to warrant his conclusion. He argues that even

if Davidson can show that sharing a prior theory is not always sufficient for communication, it does not follow that there is no such thing as a language. He writes:

> Perhaps communication does not require that *all* that a speaker and hearer 'learn in advance' match. Still, it is conceivable that a great deal of what one 'learns' must be shared in order for communication to be possible. That is, Davidson might be wrong to urge abandonment of 'the idea of a clearly defined shared structure which language-users acquire and then apply to cases.' ("Epitaphs", p. 446) This might not be the whole story, but it is an important part of it.[41]

In reply, I would like to make the same sort of argument against George that he used against Davidson: though it might be the case that much of what a speaker and hearer know before a conversation begins matches, it does not follow from this that there is a clearly defined structure that makes communication possible. It is central to Davidson's theory that much of what is learned beforehand 'matches', and also that much of what is learned beforehand continues to be shared in this way. Without these two, communication would never get off the ground. But what does not remain the same from prior to passing theory is the way a malapropism is understood. What is held in common between the speaker and the hearer, and what remains the same from the prior theory to the passing, in the interpretation of

---

[41]George, "Whose Language is it Anyway? Some Notes on Idiolects." *Philosophical Quarterly*, (July 1990), p. 285 n. 24. Emphasis in original.

one malapropism, will not remain the same for the interpretation of all malapropisms. So George's hope that there is a clearly defined, shared structure seems dashed.

For George to claim that there is a clearly defined structure that is an important part of all linguistic communication, he would need to show that there is a 'great deal' of matching, shared theory that is sufficient for interpreting malapropisms. To prove that there is such a "core" to one's language has been the goal of the pursuit in linguistics of universal rules of generative grammar. The argument that I would like to make against George is that he seems to confuse the order of quantifiers: he needs to show that there is something shared and prepared that for every successful interpretation of a malapropism remains the same from a prior theory to a passing one. Davidson, on the other hand, claims that for every successful interpretation of a malapropism, there is something that remains the same. Davidson's point is that it is impossible to specify in advance what parts have to coincide:

> The hearer must know a great deal about what to expect. But such general knowledge is hard to reduce to rules, much less to conventions or practices. (Davidson, 1984, p. 278)

I will now take up a strong objection against Davidson, articulated in the final section of Dorit Bar-On and Mark Risjord's paper "Is There Such a Thing as a Language?" (1992). The objection, as I see it, is motivated by the

appearance of an impossible demand placed on passing theories. The demand is that they be infinite in structure, and yet momentary in life-span; that they project a 'pattern at an instant'. The authors' principle objection to Davidson's argument is with the way a passing theory postulates an entire infinite language to give the interpretation of a single sentence, uttered on one occasion. They write: "A passing theory endows the momentary utterance with a structure by assigning roles to individual parts in a larger, invented whole." (Bar-On and Risjord, 1992, p. 190; emphasis added) However, the construction of a truth theory, or the empirical testing of a truth theory, happens over time. Bar-On and Risjord contend that one could never know that a passing theory is a correct theory for a language, because, to know such a thing is to know, perforce, that the T-sentence created expressly to handle the recalcitrant malaprop is entailed by an entire truth theory that optimally fits the evidence of the speaker's speech behaviour. The idea that one knows a true T-sentence to provide the correct interpretation for an utterance only if it falls out as a consequence of a truth theory that optimally fits the evidence I have called the Holistic Constraint, which is a corollary of the Holistic Thesis.

The objection is a compelling one because it takes seriously the way in which a theory of meaning is

empirically confirmed. The source of trouble for Bar-On and Risjord is this: "[Passing theories] are not responsible to multiple checkpoints; their only point of contact with the empirical reality of speech is the single utterance." (Bar-On and Risjord, 1992, p. 24) As we saw above, the strength of Davidson's truth-theoretic approach is that the empirical content of a truth theory for a language can be tested and improved by taking advantage of the fact that the theory entails law-like statements for truth-conditions in counterfactual situations. If a radical interpreter can formulate a new sentence corresponding to one of those counterfactual situations, having an extension new to the evidential base, then if that sentence is affirmed by the speaker, then we have more evidence that the theory is true. In the case of a passing theory, all but one of its T-sentences are counterfactuals. The one exception being the T-sentence that is directly connected with the particular malaprop sentence that prompted the creation of the passing theory. All others are counterfactuals *par excellence*. The spirit of a counterfactual conditional is to consider if an *a* were an F (counter-to-the-fact, for it is in fact not an F) would it be a G? The vast majority of the theorems of a passing theory are counterfactuals in the sense that they consider what the truth-conditions of a sentence that the speaker is not at the moment uttering (because the speaker is uttering at that moment only the

malaprop sentence) *would be*. However, since we do not have access to those infinite unspoken sentences, the existences of which are implied by the passing theory, we can never know whether the passing theory is well-confirmed on all its points. It is those infinite number of other sentences (the "larger, invented whole") that the objection principally focuses on. Passing theories by themselves have no empirically testable consequences, because they have no implications for the meaning of future sentences uttered. An interpreter cannot sharpen her understanding of what the malaprop means by formulating a new test sentence that contains the malaprop in the context of some other better understood words, because there is no guarantee that the deviant word or words will be used again in the same manner. Bar-On and Risjord conclude that even knowing a passing theory does not give one knowledge of an actual language, but only of a possible one. If we can't satisfy the Holistic Constraint, then we don't know whether a passing theory is well-confirmed, and thus we don't know whether the T-sentence for the malaprop sentence is true. But this is not the end of the problem. Even greater sceptical worries loom. If the passing theory can attribute idiosyncratic meaning to one particular word or expression, why can't there be many more? It seems that if there are no empirical constraints upon how one constructs a passing theory, then there is no constraint upon what theory one can 'invent' to

fill the role of a 'larger whole.' It appears that, from the starting point of a single malaprop, one can arrive at whatever language one pleases. To construct a passing theory is no more than a stipulation of a language, with no way of verifying whether it is true of some speaker's behaviour.

The question of whether a theory of truth for a language is merely stipulative or not is the central concern of Davidson's most recent publication, "The Structure and Content of Truth":

> [T]he question of whether a theory of truth is true of a given language (that is, of a speaker or a group of speakers) makes sense only if the sentences of that language have a meaning that is independent of the theory (otherwise the theory is not a theory in the ordinary sense, but a description of a possible language). ... [I]f the question can be raised ..., the language must have a life independent of the [truth] definition ... (Davidson, 1990, p. 301; emphasis added)

I will call this, following Bar-On and Risjord, the "independence requirement." As Davidson sees it, the independence requirement is a requirement to connect a theory for a language with the users of that language. The requirement asks how we know that a theory of truth for a language is correct. The thesis of Davidson's "Structure and Content of Truth" is that what is missing from any Tarski-style theory of truth is the connection with the users of the language.

Bar-On and Risjord point out, quite correctly, that there is a fundamental tie between the requirement that a theory of meaning optimally fit sentences held true by the speaker or speakers (the "holistic constraint"), and the requirement that it also fit an independent pattern. Bar-On and Risjord suggest that if we accept everything said in "A Nice Derangement of Epitaphs" then this "independence requirement" would go unfulfilled, and the truth theory would lose all empirical content. (Bar-On and Risjord, 1992, p. 187) The authors have difficulty with the way in which the passing theory is said to fit something independent of the theory and the interpreter doing the theorizing. Passing theories, because they are for the most part merely invented wholes, applicable for only an instant, they feel cannot claim to fit an independent pattern. And this does not sit well with their intuitions. As a result, Bar-On and Risjord and other critics of Davidson charge that he is advocating a "Humpty Dumpty" theory of meaning, whereby languages are completely lacking in normative elements, such that words mean whatever one wants them to mean.[42]

I will now discuss how Davidson would reply to the two objections. The first objection was that a passing theory can never be known to be true, because to be true, it has to

[42]To be fair to Bar-On and Risjord, it is only Dummett that actually levels this charge (Dummett, 1986, pp. 470 & 473). I believe that the spirit of this objection is to be found in Bar-On and Risjord (1992), and in George (1989).

make a holistic fit with an infinite amount of empirical evidence. But there is only the one sentence that a passing theory makes contact with. So, the objection goes, the Holistic Constraint goes unfulfilled. As discussed in the Introduction, Davidson's idea of a semantic norm is satisfied by any creature that behaves in patterns that on the whole allow an interpreter to construct a rationalizing, systematic theory that incorporates the abilities of prediction and explanation of speech behaviour. If what I have been arguing is correct, then the apparatus of the prior and passing theories is meant to uphold the conviction that when a speaker makes a slip of the tongue, or utters a malaprop, there still is a significant amount of semantic pattern going on that remains constant. The passing theory is merely a projection of the single idiosyncratic attribution into an existent pattern. So, Davidson is somewhat in agreement with Bar-On and Risjord when they say "the attribution of a idiosyncratic malaprop does not signal our willingness to abandon the idea that the speaker is governed by a relatively stable set of semantic norms.... In the absence of such norms, it is hard to see how a momentary utterance we try to interpret can be said to be a structured, meaningful utterance." (Bar-On and Risjord, 1992, p. 186) He agrees with them that without a pattern into which one projects the idiosyncratic malaprop sentence, it would be hard to see it as a structured sentence,

connected in logical relations with the infinity of other sentences that a language must contain. This is where the importance of the prior theory is felt. Prior theories provide the theoretical knowledge for much of the understanding of the malaprop. We understand a malaprop because the 'irregular' word (*e.g.*, "plutonic") "temporarily ...takes over the role of some other word or phrase [*e.g.*, "platonic"] <u>as treated in a prior theory</u>.... [T]he entire burden of that role, with all its implications for logical relations to other words, phrases, and sentences, [is] carried along by the passing theory." (Davidson, 1986, p. 443; emphasis added) As Dummett put it, the passing theory "massively reduplicates" the prior theory. (Dummett, 1986, p. 466) What Davidson is not tempted to do is to use the term "norm" in describing the contents of the prior theory. Any one of the "norms" written in the prior theory can be broken, at any moment. Given this, there is no longer any incentive to continue to call them norms. While it is true that a "norm" can only be broken on condition that the great majority of the other "norms" are untouched, there is no sense in which there is a "core" and a "periphery" to the language.

The upshot of all this is that a passing theory is well-confirmed based on evidence gathered over time: what someone said yesterday, and how I interpreted them, <u>is</u> evidence for the passing theory I use today, because most of

the 'building material' for the passing theory is imported wholesale from the prior theory. The passing theory cannot be invented in just any manner one pleases. If one could start from a momentary malaprop, and construct around it any language at all, as the greater whole in which to give the malaprop a semantic location and hence a meaning, then the passing theory would truly be stipulative. A stipulative definition of truth describes only a possible language. But a passing theory is postulated along the lines of things that were antecedently well understood. We do so because of our presumptions about the rationality and consistency of the agent.

The second objection was that a passing theory cannot be said to be made true by something independent of the theory. The reply is this. In general, there is something independent of a truth theory that makes it the right one or the wrong one. It is the behaviour of the speaker that is independent of the theory, and forms the evidential basis for the truth or falsity of a particular theory of truth.[43] This is something recognised by Bar-On and Risjord: "[T]he only thing that could be said to "have 'a life independent of the truth definition' would be individual utterances uttered by speakers." (Bar-On and Risjord, 1992, p. 22) Of course, Davidson does not believe that it is the case that

[43]This line of argument shares similarities with Rockney Jacobsen's unpublished manuscript "Epistemological Behaviourism and the Authority of the First-Person."

there is one unique theory which is made true by the speaker's behaviour, because of his adherence to the thesis of the indeterminacy of interpretation. As for passing theories, the test is something independent of the theory; it is still the speaker's behaviour, and whether the interpreter "got it right". Getting it right depends on a lot of factors, including what the speaker says and does next, and there are differing degrees of getting it right. The 'independence requirement' demanded by Bar-On and Risjord is fulfilled by the activities of the interpreter in that the behaviour of the speaker stands to the interpreter's theory as evidence does for any theory. When anomalous data is encountered, it is the theory that is adjusted, not the data. But such a theory is holistic, which entails that there is no way of singling out precisely which part of the theory needs changing. Furthermore, owing to the indeterminacy of interpretation, there is no unique truth theory that is made true by a speaker's behaviour. So, there is no room for realism about language.

Bar-on and Risjord object that Davidson makes out a passing theory to be some sort of constructed fiction, only useful in explaining or rationalizing a particular bit of linguistic behaviour. However, a "useful fiction" is all a language ever is for Davidson. A language is not something that is entirely independent of a theorizer. Rather, a language is something that is attributed by an interpreter

to a speaker, in order to rationalize her behaviour. As Hacking puts it, a language is not "a natural property of the person," but rather "a theoretical construct attributed to the person in the light of evidence." (Hacking, 1986, p. 453)[44] Bar-On and Risjord would then object that a "useful fiction" is not capable of supporting norms of standard, or correct usage of words. Davidson wants to de-emphasize following social standards, rules, and conventions, and emphasize the plasticity of communication, the varying degrees in which communication is more, or less, successful, and the ways we shift our communicative strategy in circumstances where conversation might be less fluid otherwise.

[44]See also Davidson (1984) pp. 239-41.

Chapter 4

"All the King's Horses"

At the end of the last chapter, Davidson was charged with having a Humpty Dumpty theory of meaning. Humpty Dumpty uses the word "glory" to mean a nice knockdown argument. Alice objects: "I don't know what you mean by 'glory'." To which Humpty Dumpty retorts: "Of course you don't -- 'til I tell you." Humpty Dumpty is claiming that he can mean by 'glory' whatever he wants. According to Dummett, the essential characteristic of a Humpty Dumpty theory of meaning is that "[a] speaker... attaches the meaning to a word by some inner mental operation." (Dummett, 1986, p. 470) Thus, Grice and Davidson get included amongst the proponents of a Humpty Dumpty theory: Grice, because his work is the *locus classicus* of the view that meaning is a sort of intending; and Davidson because of his recent adoption of, in Dummett's words, "a quasi-Gricean account of what it is to mean something by an expression." (Dummett, 1986, p. 470)

The antithesis to Humpty Dumpty's theory is Alice's. On Alice's conception, words carry meanings independently of any speaker's intentions.[45] Alice's theory makes use of what I will call the Social Thesis of language. The Social

[45]For a bit more development of this theme, see Dummett (1986) p. 473, and Hacking (1986) p. 449.

Thesis, which gains inspiration from the later Wittgenstein, finds a voice in Dummett's "The Social Character of Meaning,"[46] and is echoed in Burge, George, Bar-On and Risjord, and many others. The rough idea is this: "We hold individual speakers responsible to the norms of their linguistic community." (Bar-On and Risjord, 1992, p. 185) George, with great perspicacity, calls this idea the Communitarian view. (George, 1989, p. 288) Davidson is strongly opposed to the Communitarian conception of language, and to idea that words carry meanings independently of any speaker's intentions. He writes (rather derisively) in his most recent publication, "The Structure and Content of Truth", the following:

> There are those who are pleased to hold that the meanings of words are magically independent of the speaker's intentions; for example, that they depend on how the majority, or the best informed, or the best-born, of the community in which the speaker lives speak. ... I think this view...reveals nothing of serious philosophical interest. (Davidson, 1990, pp. 310-1)

Of the two opposing pictures of language provided by Dummett -- Alice's and Humpty Dumpty's -- Davidson certainly does not place himself amongst the supporters of the former.

Is Davidson a Humpty Dumpty theorist, as it has been defined by Dummett? We could pick flaws with Dummett's use of the phrase 'inner mental operation.' On Davidson's

---

[46]in Truth and Other Enigmas. Cambridge MA: Harvard University Press, 1978. pp. 420-30.

behalf, we could point out that on his theory it is not in virtue of an inner mental operation on the part of the speaker that words have their meaning. Rather, it is a mental operation on the part of an interpreter: the construction of a truth theory is what gives words their meanings. And this is no 'inner' operation, for it makes essential reference to a speaker, a time, and a sentence, or body of sentences.

It is true that, as with Grice's theory, there is a fundamental link in Davidson between intentions and meaning. This is what Dummett wants to bring to our attention when he discusses Davidson in connection with Humpty Dumpty. Dummett thinks that this is the source of all that is wrong with Davidson, and why Davidson is a Humpty Dumpty theorist. For Dummett, the connection between meaning and intention is nothing but a "trivial tautology, [and] not a deep truth about meaning." (Dummett, 1986, p. 472) If what I have been urging in this thesis is correct, Davidson feels exactly the same way about the connection between intentions and meaning. As we saw in chapter two, he explicitly criticised Grice for thinking that one can illuminate meaning by attending to the speaker's intentions. For Davidson, it is fruitless to attempt to solve for meaning by giving in advance content to intentions. For Davidson, there is a necessary connection between meaning and intentions, but only insofar as if one's theory of meaning attributes a

certain meaning to an expression, the theorist had better be prepared, in her wider theory of the agent's beliefs and intentions, to attribute the intention to mean what the theory of meaning says the speaker did. The intentions are constructed or attributed by the interpreter just as much as the meanings are. Whereas Grice's strategy was to discover the intentions behind speech acts, and then solve for the meanings of the utterances, Davidson does not believe that this is possible. Instead, he opts for the simultaneous attribution, by the interpreter, of both intentions and meanings. On this particular brand of instrumentalism, none of these need ever be mental events in the speaker. The speaker may not actually ever deliberately go through the inner mental operations that the theory describes. For instance, the speaker might never have the occurrent thought: "By 'monogamous' I intend to mean 'monotonous'." It is plausible that such a thought occurs with deliberate malaprops, but in the case of unintentional ones, it would seem very odd that the speaker has an intention the contents of which contain both the 'normal' and 'deviant' versions of the expression. In Chapter One, we labelled this the Representation problem: to paraphrase Dummett, we would like to say that a speaker of a malapropism believes that "The pinochle of success" means 'The pinnacle of success,' but it is a mystery how the speaker represents this knowledge to herself. Davidson's answer is to say that regardless of

whether the speaker actually forms any sort of intention or has any such kind of representation, we, as interpreters, must take the speaker as if such an intention occurred, if we are to understand the speaker at all.

Though Dummett is unsuccessful, as we just saw, in aligning the views of Davidson, Grice, and Humpty Dumpty along the lines of an 'inner mental operation' that is purportedly at work in cases of meaning, there still seems to be a point in bringing up the comparison. Davidson has indeed adopted a Gricean position. Some of the passages in "The Structure and Content of Truth" could have been lifted from some of Grice's germinal writings. For example:

> What matters to successful linguistic communication is the intention of the speaker to be interpreted in a certain way, on the one hand, and the actual interpretation of the speaker's words along the intended lines through the interpreter's recognition of the speaker's intentions, on the other. (Davidson, 1990, p. 311; emphasis added)

After which, in a footnote, Davidson acknowledges: "The influence of H.P. Grice's "Meaning" ... (1957) will be evident here." (Davidson, 1990, n. 53) Grice's 1957 paper "Meaning" comes very close to being a Humpty Dumpty theory. On a Humpty Dumpty theory, I can mean something by an utterance simply by intending to mean so; the thesis of "Meaning" is that meaning is a certain kind of intention.[47]

---

[47]To be fair, the Gricean programme, as developed by Schiffer and others, incorporated the notion that one cannot intend what is impossible, and hence avoided Humpty Dumpty-ism in

Dummett is correct in calling to our attention the similarities between Grice and Davidson. However, there are important differences in the ways they treat meaning and intentions to say that they share the same position. Dummett acknowledges that the picture that he gave of the Humpty-Dumpty theory, being characterized by an 'inner mental operation', is too crude, and easily shot down by any philosopher or linguist. But he goes on to say that Davidson's theory of meaning is nothing but a refinement of this picture. (Dummett, 1986, p. 470 & p. 473) I have tried to show how Dummett's simple picture does not apply to Davidson. Are Davidson's more sophisticated views on the nature of meaning and intention enough to distance himself sufficiently from Grice and Humpty Dumpty?

Despite Dummett's inaccurate caricature of Davidson's position, there is more than one way to skin the cat, and show how Davidson's theory can be thought of as a Humpty Dumpty theory. Dummett wants to show a positive correlation between Humpty Dumpty and Davidson, and one way of doing so is to compare their performance on the issue of linguistic error. Humpty Dumpty says that I mean whatever I intend to mean; if I believe that an utterance has a particular meaning, then it has that very meaning. If meaning is solely a matter of intending to mean something, then there can be no such thing as being mistaken about the meaning of

---

much the same way that I have shown Davidson to have done so.

a word. I could not be mistaken, because the meaning of a word is whatever I intend it to be, and in such a situation whatever seems correct to me is correct.

As for Davidson, it seems, on first inspection, that his theory also does not countenance mistakes in the way one uses language, though for different reasons. The most common reaction to a slip of the tongue or malapropism is correction. Malapropisms are *prima facie* mistakes in the language.[48] Yet Davidson is not willing to call them outright 'mistakes', because they are so readily understood. Rather, he considers malapropisms to be perfectly good cases of literal meaning, though only in a language that has a momentary lifespan, and possibly only one user. A malapropism is not 'wrong', because there is some language or other in which it is perfectly legitimate. This language is sometimes referred to as the speaker's 'instantaneous idiolect'. It is often assumed that the speaker has perfect mastery of that language.[49] In Davidson's termss, all that is required to give the malapropism the right interpretation is for the interpreter to make the necessary changes until he or she is holding a theory that is a correct one for that instantaneous idiolect. This implies that whatever unexpected turns the conversation makes, the speaker is

---

[48]A point made repeatedly by George: (1989) pp. 278-81 *passim*, and p. 288). See also Bar-On and Risjord (1992) p. 185.

[49]See George, (1989).

never wrong; the listener just has to keep up, and continually revise his or her currently held theory.

George labels Davidson's position on language a "no-error" view. (George, 1989, p. 288) He sees in Davidson the implication that there is no such thing as being in true cognitive error about facts of one's language. We will understand a Humpty Dumpty theory of meaning as implying just this. If this is truly implied by Davidson's recent philosophy of language, then we have good reason to conclude that Davidson is indeed a Humpty Dumpty theorist. The grounds for this would be independent from the ones Dummett used in the previous unsuccessful attempt to prove Davidson a Humpty Dumpty theorist by showing that Davidson's conception of linguistic intentions has the character of the mysterious 'inner mental operation' of the Humpty Dumpty theory.

The way George sees it, the Davidsonian reasoning for the conclusion that one cannot be wrong about what words mean in one's own language goes something like this: Suppose Smith calls something 'ingenuous' just in case it is ingenious. Rather than attributing to Smith a false belief about what the word means, we should instead conclude that the word 'ingenuous' in Smith's language means what 'ingenious' does in ours. While it is true that on some communal assessment of what the word means Smith was in error as to what it meant, the idea of a communal language

is not semantically significant, and hence not prior in any explanation of what someone meant by an utterance. The only sense in which Smith was 'in error' was the shallow and philosophically uninteresting notion of not aligning the way she intended to use the word with the 'common' or 'ambient' usage. (George, 1990, p. 288)

This Davidsonian argument is informed by the Principle of Charity. The overriding assumption in interpreting speakers is that they are mostly correct in the beliefs that they have about the world. In deciding on how to go about constructing a truth theory for interpreting utterances of a foreign language, a radical interpreter must make the speakers of that language come out as speaking in mostly true sentences. To paraphrase Quine, misinterpretation is more likely than speaker's stupidity. The more errors we attribute to speaker, the more we should wonder whether we have got it wrong. Attributing more and more errors about what words mean threatens the presumption that the speaker has things mostly right. At some point, it becomes more likely that the speaker is speaking a different language than the one we previously thought. An interpreter has no reason to assume that the best way of interpreting someone homophonic to the ambient speech community. There is no reason not to take a malaprop as intended to be a true sentence, and since, in the ordinary case, the changes to the truth theory that the interpreter held are easily made,

there is no reason not to say that 'Sam is ingenuous' means, in the right circumstances, 'Sam is ingenious.'

Bar-On and Risjord contend that "only commitment to an implausibly strong version of the Principle of Charity could motivate one to hold that sentences involving malapropisms should be standardly interpreted as (literally) true." (Bar-On and Risjord, 1992, p. 186) But I see no reason to call this an overly strong use of the Principle of Charity. They argued, for much of their paper, that the Principle of Charity is far too strong, and makes speakers have epistemic powers beyond what can be reasonably attributed to them. I will not engage in a protracted debate about the pros and cons of the Principle of Charity and its rivals. The 'watered-down' version of the Principle of Charity that they put forward is known as the "Principle of Humanity". Rather than maximizing true beliefs, it counsels interpreters to minimize inexplicable ones. This might be a useful maxim in some situations. In the case of a malapropism, the Principle of Humanity might provide us with the explanation that the 'mistake' occurred because the two words that were confused were phonetically similar, differing in only one syllable, *et cetera*. But such an explanation does not tell us what the malaprop means.

Bar-On and Risjord suggest that malapropisms are perfectly good examples of how the speaker can be readily understood, and yet has spoken a sentence that is literally

false; the interpreter needn't make adjustments in her theory so that it comes out a true sentence. But what they propose cannot be right. We cannot understand the notion of allegories sunning themselves on the banks of the Nile, and at the same time preserve a reasonable theory of the speaker's beliefs. For Davidson, to understand a sentence is to understand what it would be for it to be true. I know what is it for the sentence "There are allegories on the banks of the Nile" to be true; namely, that 'allegories' means alligators.

If George is right--if Davidson's theory truly is a no-error view--then Davidson's is a Humpty Dumpty theory. Unfortunately for George, Davidson explicitly denies that it is impossible for one to make a mistake in one's language. In a 1984 article entitled "First Person Authority",[50] he wrote: "The speaker can be wrong about what his own words mean." (p. 110)

There are two possible ways of reacting to this. On the one hand, one could point out that in "A Nice Derangement of Epitaphs", Davidson admitted that his most current position is at odds with some of the things that he had said in the past. This could be one of Davidson's earlier results that had to be abandoned in the face of the conclusion of the malaprop argument. On the other hand, this could still be a part of Davidson's intended position,

---

[50] *Dialectica* **38** (1984): 101-111

in which case it saves him from holding a Humpty Dumpty theory of meaning. My inclination is toward the latter. What we need, then is to show that the two papers are consistent. This requires an argument, expressed in terms of prior and passing theories, that shows how one can be wrong about the meaning of one's words. I propose to explain, using the theoretical apparatus of the "Epitaphs" paper, how it is one can be wrong about the meanings of one's words.

The theoretical apparatus introduced in the "Epitaphs" paper involves the distinction between prior and passing theories. Both the speaker and the audience form a prior and a passing theory in the course of a conversation. Most of what we have been discussing, and nearly all of Davidson's writing, focuses on the interpreter's theories. Now is the time to turn our attention to the speaker's prior and passing theories, since we are interested in how it is a <u>speaker</u> can get things wrong. Let us briefly remind ourselves again of the functions of the two theoretical devices. The speaker's prior theory describes the expectations she has about how she thinks she will be interpreted by the hearer; the speaker has to form her best picture of how she thinks her audience is 'equipped' beforehand to understand her. The speaker's passing theory is a picture of how she intends to be interpreted at each moment.

To illustrate, let us consider a case of a deliberate malaprop. Suppose that Sasha intends Hilary to understand that "polo bears" means polar bears. This would form part of the contents of Sasha's passing theory. If communication is successful, then it would also have made it into Hilary's passing theory; the passing theories would have coincided, or be shared. Suppose also that Hilary has been exposed to the way Sasha uses the word "polar", and the expression "polar bear" in the past. According to Sasha's expectations, Hilary knows that the words "polo" and "polar" are similar sounding enough such that one can make the leap of inference from one to the other: Sasha knows that Hilary will probably 'go along with it', and form the right passing theory.

If Hilary had not made the intended switch--had not kept up with Sasha's intentions--we could say that "polo bear" did not, on that occasion, mean 'polar bear': it did not mean what Sasha intended it to mean. Sasha's 'error' was not knowing that Hilary did not have the facilities to create the passing theory, or not providing enough clues. If Hilary interpreted "polo" to mean polo, then Sasha was in error about what "polo" meant. Sasha was not justified in believing that Hilary would interpret the words as intended.

This might not be enough to float our intuitions about making mistakes about what words mean. There will not always be a clear answer to the question "But what did

'polo' actually mean (on that occasion)?" Our usual understanding of what it is to be in error is that there is some 'fact of the matter'. Ordinarily, we say that someone who thinks that livid means "red" or "flushed" has got it wrong, because it actually means "pale" or "bluish gray". Does the fact that Hilary interpreted "polo" to mean polo imply that is what Sasha *actually* meant? Davidson writes:

> My characterization of successful communication leaves open a range of possibilities with respect to the question of what a speaker means by her words on occasion. ... I do not believe our standards for deciding what someone's words, as spoken on a given occasion, mean are firm enough to let us draw a sharp line between a failed intention that one's words have a certain meaning and a success at meaning accompanied by a failed intention to be interpreted as intended. (Davidson, 1990, p. 331n)

Davidson thinks that in cases where an intention to be interpreted in a certain way misfires, there is sometimes no saying exactly what the utterance meant. The salient contrast for Davidson is not between the right and wrong meaning of a word, because sometimes there is not a clear 'real' meaning in cases of misfired communication. Rather, there is a spectrum of different degrees of successful and unsuccessful communication.

Is this an adequate Anti-Humpty Dumpty argument on Davidson's part? Dummett is ambiguous on this. He writes:

> Davidson's is a version of the [Humpty Dumpty] picture. His reason for denying that, by 'glory', Humpty Dumpty could,

> in speaking to Alice, mean 'a nice knockdown argument' is ... that Humpty Dumpty knew that Alice would not understand him as meaning that. (Dummett, 1986, p. 470)

It is odd that in the first sentence just quoted, Dummett brands Davidson a Humpty Dumpty theorist, and then in the next sentence, describes the reasoning behind Davidson's rejection of the central principle of Humpty-Dumpty-ism, which is that a word means whatever one wants it to mean. Dummett still insists that Davidson is a sophisticated Humpty Dumpty theorist, despite the fact that Davidson explicitly denies the central tenet of Humpty Dumpty-ism. For Humpty Dumpty, "glory" does mean "nice knockdown argument". For Davidson, it can, but only under the right circumstances. The 'right circumstances' are determined by epistemic constraints on intentions. The epistemic constraints on what you can intend, and hence what you can mean comes down to saying that a speaker can intend for an utterance to have a particular meaning, only if the speaker is justified in thinking that the hearer will keep up with the new twist in the conversation.

We are now in a position to explain Davidson's Gricean elements. The first part of Grice's programme was a reduction of meaning in terms of the intentions held by the speaker in uttering sentences. In Davidson's system, the passing theory describes how the speaker intended to be interpreted. There is a necessary relation between how the

speaker was in fact interpreted, and her intention to be interpreted in that way. Necessarily, intended interpretation and actual interpretation must coincide in a passing theory. Whereas Grice thought that this necessary relation between intention and meaning gave one a powerful tool for discerning meaning, for Davidson, it is a vacuous truth about meaning. As for nominalism about natural languages, which is a feature of Grice's programme, Davidson's intent is to show that natural languages reduce to something more basic. Davidson sees no reason why, if a passing theory can attribute idiosyncratic meaning at a particular instant in a momentary malapropism, idiosyncratic meaning cannot be attributed in a prior theory as well.[51] He admits, though, that prior theories are what we would "most naturally describe as a natural language." (Davidson, 1986, p. 443) The less we know about someone, the more the prior theory we have of this person matches what we would call a natural language, where a natural language is some sort of generalization over a community of speakers: we start with a socially acceptable theory, and refine

---

[51] *e.g.*: "Mrs. Malaprop's theory, prior and passing, is that 'A nice derangement of epitaphs' means a nice arrangement of epithets." (Davidson, 1986, p. 443; underscoring added) I think Davidson is being incautious and imprecise with his terms. This should read: "the prior and passing theories that we, the interpreters, hold for Mrs. Malaprop's speech behaviour, is that 'A nice derangement of epitaphs' means a nice arrangement of epithets."

further.[52] But the language that a community of speakers shares has no normative force in Davidson's semantics. As Ehud Rahat puts it, Davidson acknowledges standard use or meaning in a communal language, but refuses to equate literal meaning with it.[53] The prior theory describes what a speaker's words mean, not what they mean in some communal language. Prior theories can have "all the features special to the idiolect of the speaker," writes Davidson. (Davidson, 1986, p. 443) This coincides with Grice's nominalism, in which a natural language, in the common-sense way that Finnish or Japanese are languages, is nothing more than the overlap of a great many similar, but subtly different idiolects. Davidson, I have been arguing, holds a similar view with respect to communal languages, but stops short of Grice's lower level, which is the reduction from meaning to intentions. For Grice, meaning is a kind of intention. For Davidson, by contrast, the meaning of an utterance is a matter of truth conditions given to it by a systematic theory of truth, as constructed by an interpreter. The speaker, though, can intend for the interpreter to modify the theory that she is holding, and is often successful in this intention. This link between intentions and meaning in Davidson is enough to demonstrate how he has shifted his

---

[52]Davidson, (1984) p. 153.

[53]Ehud Rahat, "Metaphors and Malapropisms: Davidson on the Limits of the Literal" in *Philosophia* v. 21, no. 3-4 (April, 1992) pp. 311-27.

position towards that of Grice. And yet, Davidson is clearly distinct from Grice, particularly in his views on the nature of belief and intention.

Bibliography

Bar-On, Dorit and Mark Risjord. "Is There Such a Thing as a Language?" in *Canadian Journal of Philosophy* **22** (1992) pp. 163-90.

Bar-On, Dorit. "On the Very Idea of a Solitary Language" in *Nous*, March 1992. v. 26, n. 1, pp. 27-45.

Bennett, Jonathan. "The Meaning Nominalist Strategy" in *Foundations of Language* **10** (1973) pp. 141-68.

_____. Linguistic Behaviour. Cambridge: Cambridge University Press, 1976.

Burge, Tyler. "Individualism and Self-Knowledge" in *Journal of Philosophy* (1988) pp. 649-65.

_____. "Individualism and the Mental" in French, Uehling, and Wettstein (eds.), *Midwest Studies in Philosophy* vol. 4. pp. 73-121.

_____. "Wherein is Language Social?" in A. George (ed.), Reflections on Chomsky. Basil Blackwell, 1989. pp. 175-91.

Cavell, Stanley. "Must We Mean What We Say?" in Must We Mean What We Say? Cambridge: Cambridge University Press, 1976. pp. 1-43.

Davidson, Donald. "Belief and the Basis of Meaning" in Donald Davidson, Inquiries into Truth and Interpretation. Oxford: Clarendon Press, 1984.

_____."A Coherence Theory of Truth and Knowledge" in E. LePore (ed.) Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson. Oxford: Basil Blackwell, 1986. pp. 159-72.

_____. "Communication and Convention" in Davidson, 1984. pp. 265- 280.

_____. "Epistemology Externalized" (ms.) 198?.

_____. "First Person Authority" in *Dialectica* **38** (1984) pp. 101-111.

_____. "Knowing One's Own Mind" Proceedings and Addresses of the American Philosophical Association **30** (1986) pp. 441-58.

_____. "The Myth of the Subjective" in M. Krausz (ed.) Relativism: Interpretation and Confrontation. Notre Dame: University of Notre Dame, 1989. pp. 159-73.

_____. "A Nice Derangement of Epitaphs" in LePore (ed.), Truth and Interpretation. pp. 433-46.

_____. "What Metaphors Mean" in Davidson, 1984. pp. 245-64.

_____. "The Structure and Content of Truth" in *Journal of Philosophy*. **87**, 1990. pp. 279-328.

_____. "Truth and Meaning" in A.P. Martinich (ed.) The Philosophy of Language, 2nd. ed. New York: Oxford University Press, 1990. pp. 79-90.

Donnellan, Keith. "Putting Humpty Dumpty Together Again" in *Philosophical Review*. **77**, 1968.

Dummett, Michael. "A Nice Derangement of Epitaphs: Some Comments on Davidson and Hacking" in LePore (ed.), Truth and Interpretation. pp. 459-76.

_____. "The Social Character of Meaning" in Truth and Other Enigmas. Cambridge: Harvard University Press, 1978. pp. 420- 430.

_____. "What is a Theory of Meaning (I)" in S. Guttenplan (ed.) Mind and Language. Oxford: Clarendon Press, 1975. pp. 97-138.

_____. "What is a Theory of Meaning (II)" in G. Evans and J. McDowell (eds.), Truth and Meaning: Essays in Semantics. Oxford: Clarendon Press, 1976. pp. 67-137.

George, Alexander. "Whose Language Is It Anyway? Some Notes on Idiolects" in *Philosophical Quarterly* **40** (1990) 275-298.

Grice, Paul. "Meaning" in *Philosophical Review* **66**, 1957. p. 377ff. Reprinted in Martinich (ed.) pp. 72-78.

_____. "Logic and Conversation" in *Syntax and Semantics* vol. 3, Peter Cole and Jerry L. Morgan, eds. New York: Academic Press, 1975. Reprinted in Martinich (ed.) pp. 149-160.

_____. "Utterer's Meaning and Intentions" in *Philosophical Review* **78** (1968) pp. 147-77.

_____. "Utterers Meaning, Sentence-Meaning, and Word-Meaning" in *Foundations of Language* 4 (1968) pp. 225-42.

Hacking, Ian. "The Parody of Conversation" in LePore (ed.) Truth and Interpretation. pp. 447-58.

Higginbotham, James. "Knowledge of Reference" in A. George (ed.), Reflections on Chomsky. Basil Blackwell, 1989. pp. 153-174.

Kripke, Saul. Wittgenstein on Rules and Private Language. Cambridge, MA: Harvard University Press, 1982. An excerpt reprinted in Martinich (ed.), pp. 475-88.

_____. "Speaker's Reference and Semantic Reference" in Peter Cole (ed.) Syntax and Semantics, Volume 9. Academic Press, 1978.

Lewis, David. Convention. Cambridge MA: Harvard University Press, 1969.

_____. "Languages and Language" in Language, Mind and Knowledge, K. Gunderson, (ed). Minneapolis: University of Minnesota Press, 1975. Reprinted in Martinich (ed.), pp. 489-508.

Peacocke, Christopher. "Truth Definitions and Actual Languages" in Evans and McDowell (eds.) Truth and Meaning: Essays in Semantics. Oxford: Clarendon Press, 1976. pp. 162-188.

Putnam, Hilary. "The Meaning of 'Meaning'" in Philosophical Papers vol. 2: Mind, Language and Reality. Cambridge: Cambridge University Press, 1975.

_____. "Reference and Understanding" in Meaning and the Moral Sciences. London: Routledge & Kegan Paul, 1978. pp. 97-119.

Rahat, Ehud. "Metaphors and Malapropisms: Davidson on the Limits of the Literal" in *Philosophia* v. 21 no. 3-4. (April, 1992) pp. 311-27.

Ramberg, Bjorn T. Donald Davidson's Philosophy of Language: An Introduction. Oxford: Basil Blackwell, 1989.

Rorty, Richard. Contingency, Irony, Solidarity. Cambridge: Cambridge University Press, 1989.

_____. "Unfamiliar Noises: Hesse and Davidson on Metaphor" in *Proceedings of the Aristotelian Society* vol. 61 (1987) pp. 283-96.

Schiffer, S. R. Meaning. Oxford: Clarendon Press, 1972.

Tarski, Alfred. "The Semantic Conception of Truth" in Martinich (ed.), pp. 48-71.

Welker, David D. "Linguistic Nominalism" *Mind* **79** (1970) 569-80

Wettstein, Howard. Has Semantics Rested on a Mistake? And Other Essays. Stanford: Stanford University Press, 1991.